

OM of: US-09-303-518d-571 to: PIR\_71.\* out\_format: pfs

Date: Jun 30, 2002 7:50 AM

About: Results were produced by the GenCode software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+np.model -DEV=x1h  
-O/cg2-1/USPTO.spool/US09303518/runat\_28062002\_142713\_4317/app.query.fasta\_1.23501  
-DB=PIR\_71 -OPMT=fastan -SUPERX=1 -GAPO=12.000 -GAPEXT=4.500  
-MINMATCH=0.100 -LOOPEXT=0.000 -OGAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-DELEX=7.000 -STAR=1 -MATRIX=blissum62 -TRANS=human40.ccd  
-LIST=100 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0  
-ALIGN=45 -MODE=LOCAL -OUTPMT=pfs -NORM=ext -HEADSIZE=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09303518.ecgn1\_1\_1092  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-303-518d-571

Query length: 894

Database: PIR\_71.\*

Database sequences: 283138

Database length: 96089334

Search time (sec): 504.980000

score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
PIR2:B81040	1469.50	2162.98	7.3e-113	298	HtrB/MSB family protein NMB1801
PIR2:D81966	1467.50	2160.03	1.1e-112	298	Probable acyltransferase NMA006
PIR2:HB3661	266.50	385.70	7.3e-14	295	Probable 2-OH-lauroyltransferase
PIR2:CB2661	258.00	372.56	3.1e-13	316	Lipid A biosynthesis lauroyl ac
PIR2:D83229	222.50	320.22	3.1e-10	312	Probable lauroyl acyltransferase
PIR2:AD0807	213.00	306.35	1.3e-09	306	Probable membrane protein YPO36
PIR2:A10441	207.50	298.22	5.3e-09	306	HtrB/MSB family protein NMB141
PIR2:D81085	202.00	293.53	1.0e-08	289	Probable acetyltransferase NMA1
PIR2:B81887	202.00	290.57	4.2e-08	306	Lipid A biosynthesis lauroyl ac
PIR2:516888	196.50	281.97	4.2e-08	306	Lipid A biosynthesis lauroyl ac
PIR2:856687	196.50	281.97	4.2e-08	306	Lipid A biosynthesis lauroyl ac
PIR2:890807	196.50	281.97	4.2e-08	306	Lipid A biosynthesis lauroyl ac
PIR2:891036	196.00	280.65	4.7e-08	328	Probable heat shock protein (im
PIR2:E75011	196.00	280.65	4.7e-08	328	Probable lauroyl acyltransferase
PIR2:D85880	192.00	280.65	4.7e-08	328	Probable heat shock protein (im
PIR2:E82847	181.50	259.81	7.3e-07	279	Lipid A biosynthesis lauroyl ac
PIR2:AD0637	175.00	246.88	2.6e-06	455	membrane-bound lytic transglyc
PIR2:E71569	168.00	229.54	3.0e-05	453	Lipid A biosynthesis lauroyl ac
PIR2:F81720	162.00	227.71	3.0e-05	290	Probable acyltransferase (EC 2.
PIR2:E71631	159.00	222.92	0.0006	343	hypothetical protein (imported)
PIR2:T42129	146.00	202.77	0.0008	1106	hypothetical protein 119.5K protein
PIR2:T42134	144.50	205.95	0.0008	290	hypothetical protein htrB (impor
PIR2:J00405	144.50	205.95	0.0008	290	probable endoglucanase - Rumunq
PIR2:828336	139.00	196.77	0.0023	315	Lipid A biosynthesis lauroyl ac
PIR2:S21323	139.00	196.77	0.0023	315	Lipid A biosynthesis lauroyl ac
PIR2:D64127	137.50	194.30	0.0030	325	Lipid A biosynthesis lauroyl ac
PIR2:F83350	137.50	194.30	0.0030	325	Lipid A biosynthesis lauroyl ac
PIR2:AE3391	137.00	190.52	0.0034	462	Lipid A biosynthesis lauroyl ac
PIR2:B72119	137.00	190.52	0.0034	462	acyltransferase - Chlamydomon
PIR2:D97552	136.50	193.16	0.0037	312	Lipid A biosynthesis lauroyl ac
PIR2:AE7772	136.50	193.16	0.0037	312	Lipid A biosynthesis lauroyl ac
PIR2:T05722	136.50	193.16	0.0037	312	Lipid A biosynthesis lauroyl ac
PIR2:CV5580	133.00	184.95	0.0079	335	adenine deaminase-related prote
PIR2:AE0521	133.00	184.95	0.0079	335	Lipid A biosynthesis lauroyl ac
PIR2:T41410	130.00	172.84	0.0126	376	exit protein - Mycobacterium s
PIR2:A55938	127.50	163.33	0.0138	1122	profilaggrin - human (fragment
PIR2:E90949	127.00	178.83	0.0234	2248	heat shock protein MsbB (import

PIR2:A85798	127.00	178.83	0.0222	323	suppressor of htrB, heat sho
PIR2:T43023	127.00	176.94	0.0226	405	ceao protein - Burkholderia
PIR2:S01359	126.50	177.97	0.0245	328	salivary glue protein Sgs-4
PIR2:A42608	125.00	175.88	0.0325	323	(Kdo)2-(lauryl)-lipid IV-3 a
PIR2:B34768	125.00	175.18	0.0327	351	Ore's protein - Orf virus (st
PIR2:T45093	123.50	169.44	0.0452	603	hypothetical protein (import
PIR2:T00062	123.00	169.68	0.0452	1571	hypothetical protein K1A04
PIR2:C39135	122.50	169.65	0.0534	437	hypothetical protein 3 (gyrB
PIR2:A84615	121.00	170.10	0.0693	318	lipid A biosynthesis lauroyl
PIR2:G83022	121.00	164.44	0.0728	625	profilaggrin - rat (fragment
PIR2:T13869	119.00	159.78	0.0829	998	probable two-component senso
PIR2:T450711	118.50	161.53	0.1161	1104	transcription factor NFX1 (
PIR2:AE57071	118.50	161.53	0.1161	569	urase (EC 3.5.1.5) alpha ch
PIR2:T07132	118.00	157.09	0.1224	323	Lipid A acyltransferase (EC
PIR2:J00984	117.50	169.11	0.1296	193	hypothetical protein 1 - Rho
PIR2:T09345	117.50	169.11	0.1296	1791	ocs-element binding factor 1
PIR2:A61183	117.00	164.35	0.1476	312	hypothetical protein K1A03
PIR2:J05280	116.50	163.40	0.1626	320	Voltage-dependent calcium c
PIR2:T05420	116.00	162.69	0.1787	319	hypothetical protein W03G1.5
PIR2:T23472	115.00	157.95	0.2223	471	hypothetical protein K08E7
PIR2:A87560	113.00	147.01	0.3485	1221	zinc finger protein (clone 1
PIR2:S36488	112.50	154.69	0.5061	688	conserved hypothetical prote
PIR2:B30843	110.50	150.92	0.5233	493	E2 protein - human papilloma
PIR2:J00689	110.50	146.71	0.5432	815	glutinin high molecular weigh
PIR2:J00059	110.50	146.71	0.5432	815	glutinin high molecular weigh
PIR2:T28682	109.50	157.38	0.5780	325	hypothetical 57.4K protein (
PIR2:D84595	109.50	157.38	0.5780	191	hypothetical protein - strep
PIR2:T44231	109.50	145.95	0.6523	748	PEAR1 A protein (imported)
PIR2:T44454	109.50	145.95	0.6523	1520	hypothetical protein 086 (1
PIR2:587040	108.50	152.70	0.6875	280	hypothetical protein T04D3.7
PIR2:C75461	108.50	148.22	0.7627	414	gene T52 protein - Trichuri
PIR2:T00273	108.00	147.00	0.8296	414	hypothetical protein - Deino
PIR2:F75311	108.00	147.00	0.8296	506	E2 protein - human papilloma
PIR2:T08176	107.50	145.54	0.9317	1520	hypothetical protein K1A05
PIR2:529693	107.00	150.22	0.9764	289	ABC transporter, Arg-binding
PIR2:828444	106.00	149.50	1.117	3938	glucosyl-1-phosphate adenylyl
PIR2:842701	106.00	149.50	1.117	264	Bassoon protein - rat
PIR2:884348	105.50	150.25	1.127	313	salivary glue protein - fruit
PIR2:526650	105.50	143.51	1.335	221	filaggrin precursor - mouse
PIR2:T83138	105.00	136.06	1.444	494	splicing factor SFRS2 - huma
PIR2:T00739	105.00	147.09	1.443	1203	hypothetical protein Vng1983
PIR2:T00204	105.00	145.88	1.444	2424	DNA-binding protein 5 - hum
PIR2:J004163	105.00	138.48	1.554	825	calcium channel B1-2 - rabb
PIR2:S02652	105.00	138.48	1.554	825	probable lipid A biosynthesi
PIR2:T06241	105.00	135.84	1.588	848	LDL receptor related protein
PIR2:T80571	104.50	146.67	1.577	1130	DNA-binding protein 585 - ra
PIR2:T40405	104.50	146.67	1.577	1130	glutinin high molecular weigh
PIR2:J004743	104.50	143.91	1.600	284	repetin - mouse
PIR2:B75310	104.50	132.66	1.777	362	histone H1 (clone TH315) - w
PIR2:T51023	104.50	127.52	1.866	395	hypothetical protein CC2597
PIR2:S62335	104.00	127.23	2.03	1513	hypothetical 44.4K protein -
	103.50	142.47	1.94	2786	hypothetical protein 086 (1
				528	fatty acid synthase (EC 2.3
				2649	conserved hypothetical prote
				393	hypothetical protein B7F21.
					I71-7 protein - fruit fly (D

seq\_name: PIR2:B81040

seq\_description\_block:

HtrB/MSB family protein NMB1801 [imported] - Neisseria meningitidis (strain MC58 ser

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence-revision 31-Mar-2000 #text-change 19-Jan-2001

C:Accession: B81040

R:RefSeq: H. K., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen,

Hi, H., H., Vamathevan, J., Gill, J., Scariato, V., Massignani, R.D., Dougherty, B.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rapunoli, R.;



```

251 GTTGGAACTGCCCCCGCTTTTTCAAAAAAGCGAAGCATCGAACA 300
|||||
84 LYLGGILLeuAlaProAlaPhePheArgLysProLysPrlIleGluThr 100
|||||
301 ATGTTCAAGCGGTACACGGCTGGGAACAGCTGACAGCGCTTTGGACA 350
|||||
101 MetPheLysAlaValHisGlyTrpGlnHisValGlnGlnAlaLeuAspLys 117
|||||
351 GGGGGAAGGCGTGTCTTCATCAGCGCGCAATCGGCACTTCAGATTGG 400
|||||
117 SHLSGLGLLeuLeuPhePheLeuThrProHisIleGlySerLysAspLeuG 134
|||||
401 GCGGAGCTPACATCAGCCAGACAGCTCCGTTCCACTGACCGCATGTAC 450
|||||
134 LysAlaGlyTrpIleSerGlnGlnLeuProPheProLeuThrAlaMetLys 150
|||||
451 AAGCGCGCAAAATCAAGCATAGCAAAATCATGCAAGGGGCGAGGCT 500
|||||
151 LysProProLysIleLysAlaIleAspLysIleMetGlnAlaGlyArgVa 167
|||||
501 GCGCGCAAGGCAAAACCGCGCCGACGATACAAAGGGTCAAAACA 550
|||||
167 ArgGlyLysGlyLysThrAlaProThrSerIleGlnGlyValLysGlnI 184
|||||
551 TCATCAAGGCGCTGCGCGCGCGGAGGAGCAATCATCTCCGCGACAC 600
|||||
184 LeileLysAlaLeuArgSerGlyLysAlaThrIleValLeuProAspHis 200
|||||
601 GTCCCTTCTCCGAGAAAGCGCG...GCGCTGTGGCGGATTTTTCGG 647
|||||
201 ValProSerProGlnGlnGlyGlyGlyValTrpValAspPhePheGln 217
|||||
648 CAACCTGTCATACACATGACATGCGGCGCAAAATTTGGCAACGTCGA 697
|||||
217 LysArgProAlaLysThrMetThrLeuAlaAlaLysLeuAlaHisValLysG 234
|||||
698 GCGGCAAAACCGCTTTTGTCTGTCGCAAGCGCTCCGCGAGCAAGG 747
|||||
234 LysAlaLysThrLeuPhePheCysCysGlnArgLeuProGlyGlnGly 250
|||||
748 TTCGTGTTCACATCCGCGCGCGCAAGGGAATTTGACGCGCAACAA 797
|||||
251 PheAspLeuHisIleArgProValGlnGlyGlyLeuAsnGlyLysAla 267
|||||
798 CCAGATGCGCGCGCTGTTCACCGCAATACCGAATTTGATAGCGCGT 847
|||||
267 HisAspAlaAlaValAlaPheAsnArgAsnAlaGlyTrpIleAlaGlyArg 284
|||||
848 TTCGAGCGGATCTGTTTATGTACACCGCTATPAAACGCGC 891
|||||
284 heProThrGlnTrpLeuPheMetLysAsnArgLysLysMetPro 298
|||||
seq_name: p1r2:H83642

```

```

seq_documentation_block:
  Probable 2-OH-lauroyltransferase PA0011 [Imported] - Pseudomonas aeruginosa (strain PA01)
  C:Species: Pseudomonas aeruginosa
  C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
  C:Accession: H83642
  R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
  adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
  Nature 406, 959-964, 2000
  A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
  A:Reference number: A82950; M01D:20437337
  A:Accession: H83642
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-295 <STO>
  A:Cross-references: GB:AE004441; GB:AE004091; NID:99945828; PIDN:AG03401.1; GSPDB:GN001
  A:Experimental source: strain PA01
  C:Genetics:
  A:Gene: PA0011

```

```

alignment_scores:
  Quality: 266.50      Length: 285
  Ratio: 1.514        Gaps: 7
  Percent Similarity: 61.754      Percent Identity: 26.667

alignment_block:
  US-09-303-518d-571 x H83642
  Align seg 1/1 to: H83642 from: 1 to: 295

55 CTGTGACGCGCCGCTGCAAAATGCTTCCTTCCTGCTGCTTCCTGCT 104
|||||
8 LeuValValGlyAlaLeuArgLeuPheAlaLeuLeuProTrpArgAlaVal 24
|||||
105 GCACACGCTGGGAACCGGCTCGACATGTGGCGCTTACTCTTAAAG 154
|||||
24 GlnGlyValGlyAlaGlyIleGlyTrpLeuMetTrpLysLeu.....P 39
|||||
155 AAGACGCGCGCGCATGCGCAATATGCGGAGCGGCTTGAACCC 204
|||||
39 roAsnArgSerArgGlyValValArgIleAsnLeuSerLysCysPhePro 55
|||||
205 .....GACACGACGCGTCAAAAGCGCTTTTCGGAACCGCA 245
|||||
56 GluLeuSerGlnThrGlnLeuGlnLysLeuValGlnIleSerLeuMetAs 72
|||||
246 ATGCGGT.....TTGCACTTGCCCGCGCTTTTCAAAACCGG 286
|||||
72 PLeuGlyArgThrLeuThrGlnSerAlaCysAlaTrpIleTrpProPro 89
|||||
287 AAGACATCGAAACAATGTTCAAAAGCGGTACACGCTGGGCAACGCGAG 336
|||||
89 LysLysSerLeuArgTrpIleArgGlnValGlnIleMetGlnValLeuGln 105
|||||
337 CAGCCTTTGGCAAGGCGCAAGGCTGCTGTTATACGCGGCACATCGG 386
|||||
106 GlnAlaLeuAlaSerGlyAspGlyLeuValGlyIleThrSerHisLeuGln 122
|||||
387 CAGCTACGAT...TTGGCGGCGCTACATACAGCGCGACGTTCCGTTCC 433
|||||
122 YAsnTrpGlnValLeuAsnHisPheTrpLysSerTrpAlaLysProLys 138
|||||
434 ACCTGACCGGCATGTCAAAAGCGCGCAAAATGCAAAAGCATACAAATC 483
|||||
139 .....IlePheTrpArgProProLysLeuLysAlaValAspLysLeu 152
|||||
484 ATGACAGCGCGGAGGCTGCGCGCAAAAGCGCAAAACCGCGCCACCGCAT 533
|||||
153 LeuLysLysGlnArgValGlnLeuGlnLysAsnArgValAlaProSerThrPr 169
|||||
534 ACAAGGGGTCAACAATCATCAAGGCGCTGCGCGCGCGAGCAACCA 583
|||||
169 OGlnGlyIleLeuSerValIleLysGlnValLysLysLysLysValG 186
|||||
584 TCATCTGCGCGCACACGCTCCCTTCGCGAGCAAGCGGCGCTGTGG 633
|||||
186 LysLeuProAlaAsp.....ProGlnProAlaArgTrpAlaGlyLeuPhe 200
|||||
634 GCGGATTTTTCGCAAACTGTCATACACATGACACTGGCGGCAAAAT 683
|||||
201 ValProLysGlnGlyThrThrAlaLeuIleSerLysPheValProGlnLe 217
|||||
684 GGCACACGTCAAAAGCGGCAAAACCGCTTTTCTGCTGCGAAGCGCTGC 733
|||||
217 LysSerArgGlyLysAlaArgGlyValIlePhePheHisAlaValArgLeuP 234
|||||
734 CCGAGGACAAGGCTTGTGTCATCCGCGCGCGCAAGGGGAGATG 783
|||||
234 roAspGlySerGlyTrpLysValIleLeuGlnAlaAlaProAlaAspMet 250
|||||
784 AAGCGCAAAAGCGCCAC...GATGCCCGCGCTTTCACCGCAATACCGA 830
|||||

```



alignment\_scores:  
 Quality: 222.50 Length: 311  
 Ratio: 1.236 Gaps: 11  
 Percent Similarity: 57.878 Percent Identity: 26.045

alignment\_block:  
 US-09-303-518D-571 x D83239 ..

Align seg 1/1 to: D83239 from: 1 to: 312

```

33 TTTGGACCGCCATCGACATCCTGTTGACGCCCT..... 68
6 PheSerSerAlaPheLeuHisProAlaTyrTrpProLeuTyrPheGlyLe 22
69 .....GCTCAATGCTCTCCCTGCTGCTGCTGCTTTC 99
22 uGlyLeuLeuTyrPheLeuValAlaGlnLeuProTyrProValLeuMetL 39
100 TGTCTGCACACCGCTGGAAACCGGCTCGACATCTGGCTTTTACCTTT 149
39 eu.....GlyArgGlyLeuGlyAlaLeuMetTyr 48
150 AAGAGACACCGCGCGCATCTGCCAATATCGCGAGCGGTTTGA 199
49 .....ArgLeuValGlySerArgArgGlnLeuAlaAlaArgAsnLeuG 63
200 A.....CCCGACACGACGAGCGGTCATA 222
63 uLeuCysPheProGlnuYsSerProAlaGlnArgGln.ArgLeuLeuLys 79
223 GCCGTTTTCGGAACGGCAAAATGGCGTTTGGAACTTCCCGCCGCTT 272
80 GlnAsnPheAlaSerSerGlyIleAlaPhePheGlnMetAlaMetSerTr 96
273 TTTCAAAAAACCGGACACATCGAAACATGTTCAAAAGCGGTACAGGCT 322
96 pTrp....TrpProYsaAlaArgLeuAlaArgLeuAlaHisIleGlnGly 112
323 GGGACACAGTGCAGAGCGTTTGGACAAAGGCGAGCGTGTTCATC 372
112 euGlnHisLeuAlaArgGlnAlaGlnAlaGlnGlyGlnGlyAlaIleLeuMet 128
373 ACCGCGCACATCGGACGTAATGTTGGCGGCGCGTACATACGCCAGCA 422
129 AlaLeuHisPheThrThrLeuGlnIleGlyAlaAlaLeuLeuGlyGln.. 144
423 GCTTCGCTCCAC...CTGACCGCCATGTACAAGCGCGCAAAATCAAG 469
145 .....ValHisThrIleAspGlyMetuYrArgGlnHisAspAsnProV 159
470 CGATGACAAATCATGACGCGGCGAGGTCGCGGCAAAAGCAAAAC 519
159 aLrPheAspArgValAlaGlnArgArgGlyArgGlnArgHisAsnLeuAspAla 175
520 GGGCCACCGGATACAAAGGCTCAAAACATCATCAAGCGCTGCGCGC 569
176 ThrAlaIleGlnuArgGlnAspValaArgSerMetLeuLysValaLeuArgG 192
570 GGGCGAGGACCAACCATCTGCGCGGACGACGTCCTTCCTCCGAGCAAG 619
192 yGlyArgAlaIleTyrArgAlaProAsp.....GlnAsp 204
620 GGGGC.....GGCGTGGCGGGAATTTTTCGCAAAACCTGCATAC 660
204 yArgAlaAlaLysGlnSerLeuPheValaProLeuPheGlyIleProAlaAla 220
661 ACCATGACACTGGCGGCAAAATTTGGACAGTCAAGGCGTGAACACCT 710
221 ThrValThrAlaThrThrLysPheAlaArgLeuGlyArgAlaArgValle 237
711 GTTTTTCGTCGGAACGCTGCCCGACGAGCAAGGCTTCGTGTGACA 760

```

```

237 uProPheThrGlnSerArgLeuAlaAspGlySerGlyTyrArgLeuThrI 254
761 TCCGCCCCCTCCAAAGGGAATTTGAACGGC...AACAAACCCACATGCC 807
254 LeHisProProLeuGlnAspPheProGlyGlnSerGlnuAlaAspCys 270
808 GCCGTGTTCACCGCAATACCGAATATTGGATATGACCGCGTTTCCGACG 857
271 LeuArgIleAsnGlnTrpValAlaGlnArgCysValaArgGlnGlnProG 287
858 GATATCTGTATTATGATACACCGCTATATAACG 888
287 nTyrLeuThrAlaHisArgArgPheLysThr 297

seq_name: p1r2:AD0807

seq_documentation_block:
probable acyltransferase (EC 2.3.1.-) [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typh
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church,
  S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
  A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
  A:Reference number: AB0502; PMID:11677608
A:Accession: AD0807
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07638.1; PID:916503625; GSPDB:GN00176
C:Genetics:
A:Gene: STY2639
C:Keywords: acyltransferase

alignment_scores:
  Quality: 213.00 Length: 293
  Ratio: 1.315 Gaps: 12
  Percent Similarity: 55.290 Percent Identity: 27.986

alignment_block:
  US-09-303-518D-571 x AD0807 ..

Align seg 1/1 to: AD0807 from: 1 to: 306

72 CAATGCTCTCCCTGCTGCTTCTGTCGACACAGCTGGAAAC 121
112 euGlnHisLeuAlaArgGlnAlaGlnAlaGlnGlyGlnGlyAlaIleLeuMet 128
31 GlnLeuProTyrProValLeuArgPheLeuGlyThrArgThrGlyLysLe 47
122 GCGTGGACATCTGGCTTTTACCTTTAAAGGAAGA..... 158
47 uAlaArg.....ProPheLeuLysArgArgGlnuSerIleA 59
159 .....CCGCGCGGCGCATGT 173
59 lAglnLysAsnIleGlnLeuCysPheProThrLeuSerArgGlnuArg 75
174 CGCAATATGCGGACGAGGCGTGTGAACCCCGACGACGACGATCAAG 223
76 GlnuLeuLeuIleAlaGlnAsnPhe.....HisSerLeuGlyMet.A 89
224 CGGTTTTCGCGGAACGCAAAATGCGTTTGAACCTTGCCTCCCGCGCTT 273
89 lAlaLeu.....LeuGlnThrGlnuMetAlaArg 97
274 TTTCAAAAAACCGGACACATCGAAACATGTTCAAAAGCGTACAGCGCTG 323
98 PheTrpProAspSerArgValArgLysTrpPheAsp...ValAspGlyLe 113

```

seq\_name: p1r2:A10441

seq\_documentation\_block:  
probable membrane protein YPO3632 [imported] - Yersinia pestis (strain C092)  
C:species: Yersinia pestis  
C:date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:accession: A10441  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltall, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-rarrago, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
NATURE 413, 523-527, 2001  
A:title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:reference number: AB0001; MUID:21470413; PMID:11586360  
A:molecule type: DNA  
A:status: preliminary  
A:residues: 1-306 <KUR>  
A:cross-references: GB:AL590842; PIDN:CAC93101.1; PID:g15981553; GSPDB:GN00175  
C:genetics:  
A:gene: IPO3632

```

Ratio: 1.289      Gaps: 9
Percent Similarity: 56.098      Percent Identity: 28.920

alignment_block:
US-09-503-518D-571 x AI0441 ..
Align seg 1/1 to: AI0441 from: 1 to: 306

```

[illegible]

277 egluLysgluilemefargalaprogluIntryleuRpreuHlsatga 294  
 878 GCTATATAAGC 888  
 294 rphelystr 297

seq\_name: p1r2:D81085

seq\_documentation\_block:

HXB/MSB family protein NMB1418 [imported] - Neisseria meningitidis (strain MC58 serogroup C)  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #extl\_change 19-Jan-2001  
 C:Accession: D81085  
 R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A. Hickey, E.R.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vignani, A.; Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; M01D:20175755  
 A:Accession: D81085  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-289 <RET>  
 A:Cross-references: GB:AE002491; GB:AE002098; NID:97226655; PIDN:AAF41779.1; PID:97226655  
 C:Genetics: Experimental source: serogroup B, strain MC58  
 A:Gene: NMB1418

alignment\_scores:

Quality:	204.00	Length:	289
Ratio:	1.207	Gaps:	12
Percent Similarity:	58.478	Percent Identity:	27.336

alignment\_block:

US-09-303-518D-571 x D81085 ..  
 Align seg 1/1 to: D81085 from: 1 to: 289

```

67 CTGCTCAAAATGCTCTCCCTGCTGCTGCTCTCTCTGTCACACGCTGGG 116
7 ValLeuTrYValLeuGlnPheLeuProPheAlaLeuLeuHlsYlLeAl 23
117 AAACCGGCTCGACATCTGCGCTTTTACCTTTAAAGACAGCGCGCGC 166
23 aaspleuThnglyleuLeuAla...TyleuLeuVallylsProAlargA 39
167 GCATC.....GTGCGCAATATGCGGACGCGCGGTTTGAACCC 204
39 rglleLylgluileasneLeuAlalysCysPheSerGluTrpserGlu 55
205 GACACGACACGCTC...AAAGCCCTTTTTCGGAACGCAAAATCGCG 251
56 LysArglysthrValLeuLysGlnHlsPheLysHlsMetAlalysLeu 72
252 TTTCGACATGCGCGCGCTTTTTCAAAAACGGAAGACATGCAACAA 301
72 tleuGluTrglYleuTrYTrpYAlarProAlglYArgleuLysSer 89
302 TGTTCAAAGCGGTACAGCGCTGGGACACGTCGACAGCGCTTTGACAG 351
89 euValArg...TyrArgasnLysHlsTyrLeuAspAspAlaLeuAla 104
352 GCGGAAGGCTGCTGTCATCAGCGCCGACATCGGACGATGACGATTTGGG 401
105 GylgluLysVallelleLeuTrYrProHlsPheThrAlarheGluMetAl 121
402 CGGAGCGTACATCAGCCAGCAGCTTCGTTSCACCTGACCGCATGTACA 451
121 aValTrAlaLeuasnGlnAspIlePro.....LeuIleSerMetLys 136

```

```

452 AGCGCCGCAAAATCAAGCGATAGACAAATCATGACGCGCGACGCTG 501
136 erHlsGlnLysasnLyslleLeuAspGlnGluLeuLysGlyleYArgasn 152
502 CGCGGCAAA.....GCAAAACCGCGCCGACCGGCGATACA 536
153 ArgTrYrHlsAsnValPheLeuIleGlyArgThr.....G 164
537 AGGGGCAAAACAAATCATCAGCGCTTCGCGCGCGGCGGCAACATCA 586
164 uLlyleuArGlnAlaLeuValLysGlnPheArGlySerSerAlarPhe 181
587 TC...CTGCGCGACCGACCTCCCTCTCTCGCGGGAAGCGCG..... 624
181 euTrYleuProAsp.....GlnAspPheGlyArgasnAsp 192
625 GCGCGTGTGGCGGATTTTTCGCAAACTGCATACCATGACACTGGC 674
193 SerValPheValAspPheGlylleGlnThrAlarHlsThrleThrGly 209
675 GCGAAATTTGCGACACGTCGCAAGCGGTGAAACCGCTTTTCTGCTCG 724
209 uSerArgTrleAlaLeuAlaAsnAlalysValleProAlaIleProV 226
725 AAGCGTGTGGCGGACGACGCAAGCGCTTCGTTGCATACCGCGCGTCA 774
226 AlArgGlnAlaAspAsnThr...ValTrleuHlsPheTrProAlarP 241
775 GGGGAATTTGAACGCGCAACAAAGCCAC...GATGCGCGCTGTTCACCG 821
242 LysSerPheProGlyleuAspAlalysAlaAspAlaGlnArgMetAsn 258
822 CAATACCGAATATGATACGCGCTTTCGACGCGAGTATGTTATGT 871
258 gPheIleGluAspArgValArgGlnHlsProGlnGlnTrYrPheTrP 275
872 ACAACGCGCTATTAACG 888
275 lslYsarPheLysThr 280

```

seq\_name: p1r2:B81857

seq\_documentation\_block:

probable acetyltransferase NMA1630 [imported] - Neisseria meningitidis (strain Z2491)  
 C:Species: Neisseria meningitidis  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #extl\_change 02-Feb-2001  
 C:Accession: B81857  
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mouton, R.; Holtroyd, S.; Jorgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
 A:Reference number: A81775; M01D:20222556  
 A:Accession: B81857  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-289 <PAR>  
 A:Cross-references: GB:AL162756; GB:AL157959; NID:97380091; PIDN:CA884858.1; PID:97380091  
 C:Genetics: Experimental source: serogroup A, strain Z2491  
 A:Gene: NMA1630

alignment\_scores:

Quality:	202.00	Length:	288
Ratio:	1.195	Gaps:	10
Percent Similarity:	58.681	Percent Identity:	26.389

alignment\_block:

US-09-303-518D-571 x B81857 ..  
 Align seg 1/1 to: B81857 from: 1 to: 289

67 CTGCTCAAAATGCTCTCCCTGCTGCTGCTCTCTGTCACACGCTGGG 116

```

7 VALLEuYrValleuGlnPheleuProheleuAlaLeuLeuHsLysLeuAl 23
117 AAACCGGCTCGACATCTGGCGCTTTACTCTTAAAGAACAGCCGCCG 166
23 aaspleuThrglyLeuAlaValLeuValLysProArgrgrArgr 40
167 GCATCGTCGCAATATGCGGAGCGGGTTGAACCCGACAGCAGACG 216
40 leGlyLutleuLeuAlaLysCysPheProGluTrpAspGlyLys 56
217 GTCAAGACCGCTT.....TTTGGGAAACGGCAATATGGCGTTT 254
57 ArGlysrThValleuLysGlnHisPheLysHsMetAlaLysLeuMe 73
255 GGAACCTGGCCCGCGCTTTTCAAAAACCGAGACATCGAACATGT 304
73 uGluTrGlyLeuTrpTrpYrAlaProAlaGlyrGlyLeuLysSer 90
305 TCAGAGCGGTACAGCGCTGGAGACAGCTGCAGCGCTTGGACAGG 354
90 alArG...TyrArGAsnLysHsTrpLeuAspAspAlaLeuAlaGly 105
355 GAAGGGTGGTGTTCATCAGCGGCACATCGGACGCTTACGATTGG 404
106 GlnLysValleuLeuLysGlnHisPheLysHsMetAlaLysLeu 122
405 ACAGCTACATCAGCGACAGCTTCGTTCCATCGTACCGCCATGTAC 454
122 ITYrAlaLeuAsnGlnAspValPro.....LeuLysMetMetYr 137
455 CGCCGAAATCAAGCATAGACAAATCATCGAGCGGAGCGGCGCG 504
137 IsGlnLysAsnLysHsLeuAspLysGlnLeuLysGlyArGAsn 153
505 GGCAGAA.....GGCAAAACCGCGCCACCGCGCATACAGAG 539
154 TyrHsAsnValPheLeuLysGlyArGThr.....GlnG 165
540 GGTCAAAACATCATCAAGCGCCCTGCGCGGCGGAGCAACCATC 588
165 YleuTrGlnAlaLeuValLysGlnPheLysSerSerAlaProPhe 182
589 ..CTGCCGACCGACGTCCTCTCCGAGAGAGCGGC.....GCG 627
182 YrLeuProAsp.....GlnAspPheGlyArGAsnAspSer 193
628 GTGTGGCGGATTTTGGCAACCTGCATACACATGACATGGCGG 677
194 ValPheValAspPheGlyLeuArGThrAlaThrLysGlyLeu 210
678 AAAATTTGGACACGTCGCAAGCGGTCGCAAAACCTTTTGTGCG 727
210 rArGlyLeuAlaLeuAlaLysValLysLeuAlaLeuProVal 227
728 GCCTGCCGCGACAGAGCTTGTGTGCATATCCGCCGCTCAAGG 777
227 rGlnLysAspAsnThr...ValTrpLeuHsPheTrpAlaLys 242
778 GAATTCAGCGGACAAAGCCGAC...GATGCCGCGCTTCAACCGCA 824
243 SerPheProSerGlnAspAlaGlnAlaAspAlaGlnArGMetAsn 259
825 TACCGAATATTGGATAGCGCGTTTCCGACGAGATGTCTTATGTA 874
259 eLleGlnGlnArGValArGlnHisProGlnLysTrpPheTrpLeu 276
875 ACCGCTATATAACG 888
276 ysArGpHeLysThr 280
seq_name: p1r2:516888

```

```

seq_documentation_block:
Lipid A biosynthesis lauroyl acyltransferase (EC 2.3.1.-) - Escherichia coli
N:Alternate names: htrb protein
C:Species: Escherichia coli
C:Date: 13-Jan-1995 #sequence,revision 13-Jan-1995 #text_change 08-Oct-1999
C:Accession: S16888; A42290; C64848
R:Karow, M.; Georgopoulos, C.
Mol. Microbiol. 5, 2285-2292, 1991
A:Title: Sequencing, mutational analysis, and transcriptional regulation of the Esche
A:Reference number: S16888; MUID:92114808
A:Accession: S16888
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <KAR>
A:Cross-references: EMBL:X61000; NID:948956; PIDN:CAN43317.1; PID:948957
A:Experimental source: strain K-12
R:Ueguchi, C.; Ito, K.
J. Bacteriol. 174, 1454-1461, 1992
A:Title: Multicopy suppression: an approach to understanding intracellular function in
A:Reference number: A42290; MUID:92165719
A:Accession: A42290
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 200-306 <UEG>
A:Cross-references: EMBL:X59939; NID:942028
A:Experimental source: strain K-12
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: C64848
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-306 <BLAT>
A:Cross-references: GB:AE00206; GB:U00096; NID:91787282; PIDN:AACT4138.1; PID:917872
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: htrb
C:Function:
A:Description: acylates the intermediate (Kdo)2-1-lipid IVA
A:Pathway: 1-lipopolysaccharide biosynthesis
A:Note: required for bacterial viability above 33 degrees C in rich media
C:Keywords: acyltransferase; 1-lipopolysaccharide biosynthesis; transmembrane protein
F:21-37/Domain: transmembrane #status predicted <TKM>

alignment_scores:
Quality: 196.50 Length: 284
Ratio: 1.268 Gaps: 8
Percent similarity: 54.577 Percent identity: 26.056

alignment_block:
US-09-303-518D-571 x S16888 ..
Align seg 1/1 to: S16888 from: 1 to: 306
103 CTGACAGCTGGGAAACCGGCTCGACATCTGCGCTTTTACTTTAA 152
227 rGlnLysAspAsnThr...ValTrpLeuHsPheTrpAlaLys 242
37 lIeYrArGleuGlyCysGlyLeuGlyLysLeuAlaLeuArGpHeMety 53
153 G..... 153
1 sarGArGAlaLysIleValHsArGAsnLeuGlnLeuCysPheProGlu 70
53 .. 70
154 .. 86
70 eLsErGlnGlnArGValArGlnHisProGlnLysTrpPheTrpLeu 276
193 GGTTCAGACCCGACAGCAGAGCGGTCAAGCCGTTTGGGAAACGCG 242
87 GlyMetGlyLeu..... 90

```











C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_rev1sion 20-Aug-2000 #ext\_change 20-Aug-2000  
C:Accession: E82847  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82151; MIMD:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: E82847  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-279 <SIM>  
A:Cross-references: GB:AEO03864; GB:AEO03849; MID:9104879; PIDN:AAE82917.1; GSPDB:GN001  
R:Experimental source: strain 945C  
R:Stimpson, A.J.G.; Reilach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Britton, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, A  
das-Neto, E.; Docena, C.; El-Porri, H.; Facinanci, A.P.; Ferreira, A.J.S.  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tashiro, M.H.; Vailada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0104

## alignment\_scores:

Quality:	193.00	Length:	270
Ratio:	1.255	Gaps:	10
Percent Similarity:	56.667	Percent Identity:	27.407

## alignment\_block:

US-09-303-518D-571 x E82847

Align seg 1/1 to: E82847 from: 1 to: 279

```
112 CTGGGAACCGGCTGGACATCTGGCTTTACCTTTAAAGAACCG 161
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
15 LcnglYArgValIleGlYAlaMetAlaLeuArgValIGlYThrArg 31
162 CGGGGCGATCGTGGCAATATCGGCGAGGCGGTTGAACCCGACACGC 211
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
31 gCysAlaAlaGluValAsnLeuArgLeuCySPherProThrHisAspGlu 48
212 AGACGCGTCAAGCGCGTTTTCGGGAAGCGCAAAATGC.....GCT 252
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
48 lATrArgLysLysLeuHisLysGluAsnPhrAlaLeuGluValGly 64
253 TTG...GAACCTGGCCCGCGCTTTTCAAAAACCGGAAGACATGGAAC 299
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
65 LeuPhgLeuPheAlaArgAlaTrpTrp.....GlySerIleAspVal 78
300 AATGTTCAACCGGTACAC.....GGCTGGGAACAGATGACATGCAAC 343
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
78 lLcnglArgLysLysLysValGluGluLysLysLysLysLysLys 95
344 TGGACAGCGCGAAGCGCTGCTTTCATCAAGCGGACATGCGACGATAC 393
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
95 lnaTgGluGluArgGlyValLeuLeuValSerGlyHisPheMetThrLeu 111
394 GATTGGCGGCGACGCTACATCAGCAGACATTCGCTGACATGACGCG 443
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
112 GluLeuGluArgLysLysLysLysLysLysLysLysLysLysLys 126
444 CATATCAAGCGCGCAAAATCAAGGATAGCAAAATCATCATGACGCG 493
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
126 yMeTTrArgArgHisArgAsnProValLeuGluTrpAlaIleLysArg 143
```

seq\_name: p1r2:AD0637

## seq\_documentation\_block:

lipid A biosynthesis [auroy] acyltransferase (EC 2.3.1.-) (heat shock protein) [import  
C:Species: Salmonella enterica subsp. enterica serovar Typh  
A:Note: this species has also been called Salmonella typh  
C:Date: 09-Nov-2001 #sequence\_rev1sion 09-Nov-2001 #ext\_change 27-Nov-2001  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AD0637  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-306 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD08279.1; PID:916502326; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY1192  
C:Keywords: acyltransferase

## alignment\_scores:

Quality:	181.50	Length:	276
Ratio: <td>1.134</td> <td>Gaps:<td>12</td></td>	1.134	Gaps: <td>12</td>	12
Percent Similarity: <td>57.971</td> <td>Percent Identity:<td>26.087</td></td>	57.971	Percent Identity: <td>26.087</td>	26.087

## alignment\_block:

US-09-303-518D-571 x AD0637

Align seg 1/1 to: AD0637 from: 1 to: 306

103 CTGCACACGCTGGGAACCGGCTGGACATCTGGCTTTTACCTTTTAA 152



```

520 GGGCCACCGGATACAGGGGTCAACAAATATCAAGAGCCCTGCGCC 569
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
194 ValProGln...AsnAlaIleAsnGlnAlaLeuArgAlaLeuHisr 209
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
570 GGGCGAGCAACATCATCTGCCGACACAGCTCTTCCGAGGAG 619
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
209 gElgyluValValGlyIleValGlyAspIleValLeuLeuSerSergIut 226
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
620 GGGGGGGCGGTGTGGCGGATTTTTCGGCAAACTGCATACACCATGACA 669
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
226 yfser.....TyrProLeuPheGlySergIleAlaPheThrThr 239
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
670 CTGGGGGCAAAATTTGGCACACGCTCAAGAGCGTGAACCTGTTTTCG 719
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
240 SerProAlaLeuLeuAlaIleValThrLysThrLysThrValIleAlaValAl 256
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
720 CTGGCAACGCGTCCCGCAGCAGCAAGGTTGTGTGTGCACATCCGCCCG 769
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
256 aIleTyrAlrGlySproAsnGlyAsnTyrLeuValVal..... 268
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
770 TCCAAGGGGAATTGAACGGCAACAAGCCACGATCCCGCGTGTTCAC 819
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
269 .....ProSerLysAlaPheHis 274
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
820 CGCAATACCGGAATTTGATACG..... 843
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
275 AlaAsnThrGluLeuSerIleArgLysSerThrGluGlnLeuMetAsp 291
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
844 .....CGTTT.....CCGACGAGTATC 862
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
291 gLeuMetArgPheLeuGlnLysGlyIleThrGlySlysProGluGlnIPL 308
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
308 eutIleuHisLysArgTyrLys 315

```

seq\_name: p1r2:164053

seq\_documentation\_block:

membrane-bound lytic transglycosylase homolog - Haemophilus influenzae (strain Rd [KW20])  
 C:Species: Haemophilus influenzae  
 C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
 C:Accession: 164053  
 R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; D.M.; Brandon, R.C.; Scott, J.; Shirley, R.; Liu, L.I.; Glodet, A.; Kelley, J.M.; Weidman, J. Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Reference: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Accession: 164053  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-318 <TIGR>  
 A:Cross-references: GB:U32705; GB:U42023; NID:G1573156; PID:AC21868.1; PID:G1573159; T

alignment\_scores:

Quality: 168.00 Length: 265  
 Ratio: 1.105 Gaps: 10  
 Percent Similarity: 57.358 Percent Identity: 24.151

alignment\_block:

US-09-303-518d-571 x 164053 ..

Align seg 1/1 to: 164053 from: 1 to: 318

```

112 CTGGGAACCGGCTCGACATCTGGCGTTTACCTTTAAAGAAGACCG 161
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
57 LeuGlyIleTrrIleGlyHisLysAla.....LysLysGlnAr 69
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
162 GCGCGCATCGTCCCATATGCGGAGCG.....GTTTGA 199
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

seq\_name: p1r2:F81720

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase, probable TC0278 [imported] - Chlamydia  
 C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn  
 C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: F81720  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39  
 A:Reference number: AB1500; MUID:20150255  
 A:Accession: F81720  
 A:Status: preliminary  
 A:Molecule type: DNA

```

69 gThrArgAlaGlnThrAsnLeuGlnTyrCysPheProHisTrrThrGlu 86
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
200 ACCCGACACGACGCGTCAACACCGCTTTTGGGAAACGCAAAATGC 249
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86 IngIaValGluGlnValIleIleAspLysMetPheAlaValAlaIleVal 102
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
250 GATTGGAACTTGGCCCGCGTTTTCAAAAAACCGGACACATCGAAAC 299
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
103 MetPheGlyIleGlyGlnIleAlaIleArgSerLysLysHisLeuGln 119
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
300 AATGTTCAAGCGGTACACGCTGGGACACGTCGACAGCGCTTTGAGA 349
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
119 sArgSerGluPheIle...GlyLeuGlnHisIleGlnIleAlaLysAla 135
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
350 AGGCGAAGGCGTCTGTTTCATCAGCGCACATCGGACGCTACGATTGG 399
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
135 IngLysAsnIleIleLeuMetValProHisGlyTrrPalaIleAspAla 151
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
400 GCGGACGCTCATC...AGCCAGACGTTCCGTTCCACCTGACCGCAT 446
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
152 SerGlyIleIleLeuHisThrGlnGlyMetPro.....MetThrSern 166
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
447 GTACAGCGCGCGCAATATCAAGCGTACACAAATATCAGCGGCGCA 496
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
166 tTyrAsnProHisAlrGAsnProLeuValAspTrrPleuTrrIleThrA 183
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
497 GGGTGGCGGCAAGGCAAAACCGCGCCACCGCATACAGAGGGTCAA 546
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
183 rGlnArGrrPheGlyGlyLysMetHisAlaArgGln...AsnGlyIleLys 198
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
547 CAATATCATAGGCCCTGCGCGCGCGGACGACCATCATCTCCGCCGA 596
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
199 ProPheLeuSerHisValAlrGlySglYleuMetGlyTyrTyrLeuProAs 215
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
597 CCAGCTCCCTCTCCGAGAGAGCGCGCGCTGTGGCGGATTTTCG 646
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
215 pGlnAspPheGlyAlaGln.....SerValPheValAspPheG 230
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
647 GCAAACTGTCATACACCATGACACTGCGCGCAAAATTTGCACAGCTC 693
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
230 lYThrTyrLysAlaThrLeuProGlyLeuAsnLysMetAlaLysLeuSer 246
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
694 AAGCGCTGAACACCTGTTTTCGTCGAGACGCGCTCCCGCAGGACA 743
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
247 LysAlaValValIlePheMetPheProArGlyrAsnAlaGlnThrGlyL 263
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
744 AGGCTGCTGTGCACATCCGCCCTGTCAGAGGAAATTGAACGCGACA 793
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
263 s...TyrGluMetGluIleHisProAlaMet...AsnLeuSerAsp 278
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
794 AAGCCACGATGCGCGGTTCACACCGCAATACCGAATTTGGATTCGC 843
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
278 roGlnGlnSerAlaArgAlaMetAsnGlnGlnIleGluSerPheValThr 294
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
844 CGTTTCCGACGACATCTGTTTATGTACACCGCTATTAACG 888
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
295 ProAlaProGluGlnTyrValTrrPleLeuGlnLeuLeuArGrrThr 309
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

A:Residues: 1-453 <TEF>  
 A:Cross-references: GB:AE002295; GB:AE002160; NID:97190314; PIDN:AAF39146.1; PID:97190314  
 A:Experimental source: strain N199 (MOPn)  
 C:Genetics:  
 A:Gene: TC0278

alignment\_scores:  
 Quality: 162.00 Length: 342  
 Ratio: 0.970 Gaps: 13  
 Percent Similarity: 48.830 Percent Identity: 23.977

alignment\_block:  
 US-09-303-518D-571 x F81720 ..

Align seg 1/1 to: F81720 from: 1 to: 453

```

22 CTGTTTCCCTTGGGACCGCCATGCATCTCTTG..... 60
   ||||| ||||| ||||| ||||| ||||| |||||
2 LeuphelysmetleuargserglylylvalleuvalasphisleuVa 18
   .....ACGCGCTGCTCAATGCTCTCCCTGCTGCTGCTT 97
61 ..... ||||| ||||| ||||| ||||| |||||
18 lTyrglyleuglyleuglylleuthrilleuargyleuProargS 35
   ||||| ||||| ||||| ||||| ||||| |||||
98 CCTGCTGCACACGCTGGGAAACCGCTCGACATCTGGCGTTTACTT 147
   ||||| ||||| ||||| ||||| ||||| |||||
35 erSerleuglnleupheglylylserglylylthrThrillepHeTyVal 51
   .....TTAAGGAAACCGCGCGCGCATGCTGCGCAATATGCGG..... 186
148 TTAAGGAAACCGCGCGCGCATGCTGCGCAATATGCGG..... 186
   ..... ||||| ||||| ||||| ||||| |||||
52 lIleerSperhearglysthrAlaleuthrAsnleuAlaleuAlaphePr 68
   ..... ||||| ||||| ||||| ||||| |||||
187 ..... ||||| ||||| ||||| ||||| |||||
68 oglnlyserpherthrlyuargtyrGlnllealaleuLys...SerValG 84
   ..... ||||| ||||| ||||| ||||| |||||
212 AGACGGTAAACCGCTTTTGGGAAACCGGCAAAA..... 246
   ||||| ||||| ||||| ||||| ||||| |||||
84 lnglnVallleIleThrheIleuglnleuAlathrValasplysPheAla 100
   ..... ||||| ||||| ||||| ||||| |||||
247 ..... ||||| ||||| ||||| ||||| |||||
101 lYshIleleaspglnlleIleThrIlealaserSerGlnuAspIleProgl 117
   ..... ||||| ||||| ||||| ||||| |||||
268 .GCGTTTTTCAAAAAACCGGAAACATGCAAAATGTTCAAAAGCGTAC 316
   ..... ||||| ||||| ||||| ||||| |||||
117 uGlyPhePhe.....ProglnglnuValserSerGlnnglnleu..... 130
   ..... ||||| ||||| ||||| ||||| |||||
317 ACGGCTGGGAACAGTGGACGCTTTGGACAAAGGCGGAGGCTGCTG 366
   ..... ||||| ||||| ||||| ||||| |||||
131 ..... AsnAsnpherheSerargyleuAspGlnnglnlyalale 144
   ..... ||||| ||||| ||||| ||||| |||||
367 TTCATCACCCCGACATCGGACGTAAGATTGGCGGACGCTACATCAG 416
   ..... ||||| ||||| ||||| ||||| |||||
145 leupheCyglyHieGlnAlaAsntrpGlnleuProPheleuTyrlleH 161
   ..... ||||| ||||| ||||| ||||| |||||
417 CCAAGCAGCTTCGCTTCACCTGACCGCCATGTACAAGCCGCGGAAATCA 466
   ..... ||||| ||||| ||||| ||||| |||||
161 rlysaTygTyPProgly.....leuAlaPheAlaLysProValLysAsnR 176
   ..... ||||| ||||| ||||| ||||| |||||
467 AAGCGATAGCAAAATCATGACAGCGGCGGACGCTGGCGGCAAGGCAAA 516
   ..... ||||| ||||| ||||| ||||| |||||
176 roArgleuAsnArglylleIleSerleuArgLysPheGlnnglyLys 192
   ..... ||||| ||||| ||||| ||||| |||||
517 ACCGCGCCACCGGCAATACAGGGGTCAAAACAATATCAAGGCGCTGG 566
   ..... ||||| ||||| ||||| ||||| |||||
193 lIleValIleProProln...AsnAlaIleAsnGlnAlaLeuArgAlaLeuH 208
   ..... ||||| ||||| ||||| ||||| |||||
567 CGGCGGAGGCAACCATCATCTCTGCGCGACGATCCCTTCGCGCAG 616
   ..... ||||| ||||| ||||| ||||| |||||
208 slysglylnuValValglylleValGlyAspGlnValleuLeuSerSerG 225

```

```

617 AAGCGGCGCGCTGTGGCGGATTTTTCGCAAACTGCATACACCAAG 666
   ..... ||||| ||||| ||||| ||||| |||||
225 lnyrSer.....TyrProleuPheglySerGlnAlaPheThr 238
   ..... ||||| ||||| ||||| ||||| |||||
667 ACACTGGCGCAAAATTTGGACAGTCAAAAGCGTGAAAACCTGTTT 716
   ..... ||||| ||||| ||||| ||||| |||||
239 ThrSerProAlaLeuAlaTyrlsThrArgLysProValIleAlaVa 255
   ..... ||||| ||||| ||||| ||||| |||||
717 CTGCTGGCAACCGCTGCGCGGACGCAAGGCTTCGTTGTCATCCGCC 766
   ..... ||||| ||||| ||||| ||||| |||||
255 lAlaIleTyrglnProAsnGlyAsnTyrlleVal..... 268
   ..... ||||| ||||| ||||| ||||| |||||
767 CCGTCAAGGCGGAATGTGAACGCAAAAGCCACGATGCCCGCTGTC 816
   ..... ||||| ||||| ||||| ||||| |||||
269 .....ProSerLysAlaPhe 273
   ..... ||||| ||||| ||||| ||||| |||||
817 AACCGCATACCGAATATGGATACG..... 843
   ..... ||||| ||||| ||||| ||||| |||||
274 TyrlaAsnThrGlnleuSerIleArgLysThrGlnleuMetas 290
   ..... ||||| ||||| ||||| ||||| |||||
844 .....CGTPT.....CCGACGCACT 859
   ..... ||||| ||||| ||||| ||||| |||||
290 pLysleuMetArgPheleuGlnLysGlylleAlaCysLysProglnglnT 307
   ..... ||||| ||||| ||||| ||||| |||||
860 ATCTGTTATGTACACCGCTATATA 885
   ..... ||||| ||||| ||||| ||||| |||||
307 rpleuThrpleuHieLysArgTrpLys 315
   ..... ||||| ||||| ||||| ||||| |||||
seq_name: pLr:E71631

```

seq\_documentation\_block:

lipid A biosynthesis lauroyl acyltransferase (htrb) RP718 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C:Date: 21-Nov-1998 #sequence,revision 21-Nov-1998 #text,change 03-Nov-2000  
 R:Accession: E71631  
 R:Andersson, S.G.E.; Zomrodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark  
 Nature 396, 133-140, 1998  
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A71650; MUID:99039499  
 A:Accession: E71631  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-290 <AND>  
 A:Cross-references: GB:AJ235273; GB:AJ235269; NID:93861237; PIDN:CAA15149.1; PID:e134  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: htrb; RP718

alignment\_scores:  
 Quality: 159.00 Length: 205  
 Ratio: 1.214 Gaps: 9  
 Percent Similarity: 63.902 Percent Identity: 21.951

alignment\_block:  
 US-09-303-518D-571 x E71631 ..

Align seg 1/1 to: E71631 from: 1 to: 290

```

289 GACATCGAACAATGTTCAAGCGGTACACGCTGGGAAACGTCGACGA 338
   ..... ||||| ||||| ||||| ||||| |||||
95 GluLeuGlnSerArgIleGlnIleIle...GlyLysGlnSnIleLysly 110
   ..... ||||| ||||| ||||| ||||| |||||
339 GCGTTTGACAAAGGCGGAGGCTGCTGTATCATCACCGCCACATCGCA 388
   ..... ||||| ||||| ||||| ||||| |||||
110 sleuGlnuSp.....GlnProPheleuLeuPheSerGlyHisPheAla 125
   ..... ||||| ||||| ||||| ||||| |||||
389 GCTACGATTGCGGCGAGCATACATCAGCCAGCAGCTTCGTTCCACCTG 438
   ..... ||||| ||||| ||||| ||||| |||||
125 snrPaspIleSerleuLysValleuLysLysTyTYrTYPro...LysVal 140
   ..... ||||| ||||| ||||| ||||| |||||
439 ACCGCAATGTACAAGCCCGCAAAATCAAGCGATAGCAAAATCATGCA 488
   ..... ||||| ||||| ||||| ||||| |||||

```



```

141 AlaValIleTyrArgLysAlaAsnProTyrValAsnLysLeuValAs 157
489 GGGGGGAGGTCGGCGGCAAAAGCGCCACCGCATATACAG 538
157 nclusterrargLysAspLysLeuArgLeuIleProLysLysArgLug 174
539 GGGTCAACCAATCATCAAGCGCGCGGCGGAGGAGGACCATATC 588
174 LylleArgSerLeuValArgAlaIleLysGluLysLysLysLeuMet 190
589 CTGGCCGACACGCTCTTCCCGAGAGAGCGCGCTGTGGCGGA 638
191 LeuValAspGln.....LysMetAsnAspGlyIleGluValPr 203
639 TTTTTCGCAACCTCATACCATGACATGCGCGCAAAATTTGGCAC 688
203 ophleucllyArgProSerMetThrAlaSerSerIleAlaLysIleAla 219
689 ACGTCAAGGCGTGAACCCCTGTTTCTGTCGCGAGCGCCG...CCC 735
220 .....LeuGlnTyrLysTyrProIleIleProCysGlnIleIleArgThr 234
736 GACGAGACAGGCTTCGTCGACATCCGCGCC...GTCCAGGAGGAT 782
235 LysGlySerTyrPheLysValIleValHisProGlnLeuLysPheGlu 251
783 GAACGGCAACAAGCCGACGATGCC.....GCCGTTCACACCGCAAT 825
251 ntttGlyAspAsnLysValAspCysTyrAsnIleMetPheAsnIleAsnG 268
826 .....ACCGAATATGTGATACGCGCTTTCGACGACGATATCTGTTATG 870
268 InleuGluGlyIleTyrValLysGlnAsnProAlaGlnIlePheIlePhe 284
871 TACACCGCTATATA 885
285 HisAsnArgTyrLys 289

```

seq\_name: p1r2:T42129

seq\_documentation\_block:  
 C:Probable acyltransferase (EC 2.3.1.-) - Escherichia coli plasmid p0157  
 C:Species: Escherichia coli  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Nov-2000  
 C:Accession: T42129; T00321  
 R:Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.  
 Nucleic Acids Res. 26, 4196-4204, 1998  
 A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli serotype O157:H7  
 A:Reference number: 222068; PMID:98391744  
 A:Accession: T42129  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-343 <BUR>  
 A:Cross-references: EMBL:AF074613; PIDD:AAC70097.1  
 A:Experimental source: strain EDL933; serotype O157:H7  
 S.: Shingawa, H.  
 DNA Res. 5, 1-9, 1998  
 A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic E. coli O157:H7  
 A:Reference number: T4127; PMID:98290540  
 A:Accession: T00321  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 12-343 <MAK>  
 A:Cross-references: EMBL:AB011549; NID:94589740; PIDD:BAA11840.1; PIDD:93337081  
 A:Experimental source: strain EHEC O157:H7, substrain RMD 050952  
 C:Genetics:  
 A:Genome: plasmid p0157  
 A:Note: L7029  
 C:Keywords: acyltransferase

alignment\_scores:

Quality: 156.50 Length: 273  
 Ratio: 1.057 Gaps: 10  
 Percent similarity: 54.212 Percent identity: 23.810

alignment\_block:

US-09-303-518D-571 x T42129 ..

Align seg 1/1 to: T42129 from: 1 to: 343

```

121 CGGCTCGACATCTGGCTTTTACCTTTTAAAGAAAGCGCGCAT 170
80 LysLeuGlyMetLeuValGluArgLeuGlyLysSerAlaArgGlnArgAl 96
171 CGTGGCAATATGCGG.....CAGCGGGCTTGAACCCGACGA 208
96 AleuIleAsnLeuSerLeuLysPheProGluTyrSerAspLysGluLysG 113
209 CGCAGACGTCMAACCGCTTTTTCGGAACGCGCAAAATGCGGTTGGA 258
113 LysAsnIleValAspAlaMetPheAlaThrAlaSerMetAlaValValLeu 129
259 CTGGCCCGCCCGCTTTTCAAAAACCGGAAGACATCGAACAATGTGCA 308
130 MetAlaIleLeuAlaLeuSerGlyProAspLysIleSerHisArgIleAr 146
309 AGCGGTACACGCGTGGGACACGTCGACGAGCGCTTGGCAAGGCGGAG 358
146 g...TPasnGlyLeuGlnIleValGluLysMetAlaGlnAsnGlnL 162
359 GCGTGGCTGTCATCACGCGGACATCGGACGAGCTACGATTCG..... 399
162 yValIleAlaPheLeuValProHisAlaTyrGlyValAspIleProAlaMet 178
400 .....GGCGACGCTACATCAAGCCAGCGAGCTTCGTCACCT 437
179 LeuMetAlaAlaSerGlyArg.....LysMe 187
438 GACCGCATGTACAAAGCCCGCAAAATCAAGATACAAATCATGCG 487
187 ValAlaMetPheHisAsnGlnArgAsnProValValAspTyrAlaIlePr 204
488 AGCGGCGACAGGTCGCGCGCAAGGCAAAAGCGCGCCACCGGATGCA 537
204 snserValArgArgPheGlyLysLysLeuHisAlaArgAsn...Asp 219
538 GGGGTCAACAATATCATCAAGCGCTGCGCGCGGCGGCAACCATCAT 587
220 GlyIleAlaSerPheValArgSerValArgGlnGlyTyrIlePheLys 236
588 COTGGCC.....GACACGTCCTTCTCCGAGGAGGCGGCGGCTGT 631
236 rLeuProAspGlnAspHisGlyProGluPheSerGlu.....P 249
632 GGGGCGATTTTTCGCAAACTGCATACACATGACACTGGCGGCAAA 681
249 heAlaAspPhePheAlaThrTyrLysAlaPheProValIleGlyArg 265
682 TTGGCAGCGTCAAAAGCGGTGAAA.....ACCGTGTTCCTGCTGCGA 725
266 LeuSerArgLysSerGlyAlaArgIleIleProLeuPhe..... 278
726 ACGCTGCCCGACGACAA.....GGCTTGTGTGACATCGCGCGCG 769
279 .ProValTyrAspGlyLysThrHisLysLeuThrIleHisValSerPro 295
770 TCCAGGCGAATTGACGCGCAACAAAGCCGACATCGCGCGGTTCAC 819
295 roLeuAlaIleArgGlnLysSerAspAlaHisIleAlaArgGlnIleAsn 311
820 CGCAATACCGAATATGATACGCGCTTTCGAGCGCATGCTGTAT 869
312 GluValValGluAsnPheValArgProHisProGluGlnTyrThrIleP 328

```

870 GTACACCGCTATATAACG 888  
 : : : : :  
 328 eLeuylsLeuLeuylsThr 334

seq\_name: p1r2:T45134

seq\_documentation\_block:  
 hypothetical protein [imported] - Microbacterium ammoniaphilum (fragment)

C:Species: Microbacterium ammoniaphilum  
 C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000

C:Accession: T45134

R:Strleel, H.M.; Seeber, S.; Jarsch, M.; Kessler, C.  
 Gene 172, 41-46, 1996

A:Title: Cloning and characterization of the Mami restriction-modification system from

A:Reference number: 222923; MUID:96257250

A:Accession: T45134

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-529 <STR>

A:Cross-references: EMBL:X79027; NID:9984667; PIDN:CAA55649.1; PID:g1679831

A:Experimental source: ATCC 15354

alignment\_scores:  
 Quality: 146.00 Length: 320  
 Ratio: 1.090 Gaps: 19  
 Percent Similarity: 41.875 Percent Identity: 27.188

alignment\_block:  
 US-09-303-518D-571 x T45134 ..

Align seg 1/1 to: T45134 from: 1 to: 529

```

30 CCCCTTTCGACACCGCATGACATCTGTTGACCGCCCTGCTCAATGCG 79
   ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
74 ProAlaGlnArgGlnValHisProLeuGlnArgProHisGlnLys 90
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
80 T.....CTCCCTGCTGCGCTTCTGCTGCACACGCTG 114
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
90 OGlyAlaArgGlyGlyLeuPro...ValArgAspArgAlaGlyArg 106
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
115 GAAA.....CCGCTCGGACATCT..... 134
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
106 LysArgGlnArgProAlaArgProAlaAlaArgAspAlaArgHisLeu 122
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
135 .....GGCGTTTACCTTTT...AAAGAAAGACCGCGCGCATGCTG 175
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
123 ArgLeuGlyAlaHisProAlaArgArgGlyValLeuArgArgHisAla 139
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
176 CCAATATGCGGAGGCG.....GGTTTGAACCCGCA..... 206
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
139 GHisArgAlaArgGlyGlyArgGlyArgGlyProArgGlnGlnValPro 156
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
207 .....CAGCGACGCGTCAAGCGCTTTTTCGGAACGGCAAAATGCG 251
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
156 rGlnHisProArgGlyArgArgAspArgAlaGly..... 167
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
252 TTTGCAACTGCGCGCGCTTTTCAAAAACCGAAGACATCGAACA 301
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
168 .....ArgProGlyLeuHisArgArgArg 175
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
302 TGTTCAAAGCGGTACACGGCTGGAACACGTGCACGAGCTTTGGACAG 351
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
175 gATgATgAlaArgArgArgGlnGlyArgProGlnValArgHisGly 192
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
352 GCGGAAGGCGTCTGTTTCATCAGCGCGCACATCGGCACTGAGATTGG 401
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
192 spGln.....HisArgAlaAspProArgArgProArgProArgAsp 204
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
402 CGGACGCTATATGACGACGAGCTTCGTTCCACCTGACCGGCAATGTA 451
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
205 ArgAlaGlyHisHisProLeuArg.....G1 213
```

```

452 AGCGCGCGAAATCAAGCGATGACAAATCATGCA...GGCGGGCAGG 498
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
213 uGlyGlyGlnGlyGlnGlnValArgProLeuArgAlaArgGlyGlnG 230
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
499 GT..... 500
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
230 LysGlyProGlyGlyProAlaGlyArgHisProAlaLeuGlyAspGly 246
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
501 .....GGCGGCGAAAGCAAAACCGCGCGCCACCGG.....CA 532
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
247 ArgProArgAlaArgGlnGlnGlyAlaArgAlaHisArgGlnAlaLeu 263
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
533 TACAAAGGCGTCAACCAATCATCAAGGCGCGCGCGGCGGAGGCAACC 582
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
263 sLeuArgLeuGlnArgArg.....ArgGlyArgAlaAsp 275
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
583 ATCATCTGCGCCGACACGCTGCTTCCGAGAGAGCGCGCGCGTGG 632
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
275 LysArgAlaGlnGlyArg.....AlaArgAlaArgAlaArgAla 288
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
633 GCGGATTTTTCGGCAACCTGCATACCATGACATGCGCGCGCAAAAT 682
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
289 GlyGly.....GlyGly..... 292
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
683 TGGCACACGTCAAAGCGTGAACCTGTTTCTGCTCGCAACGCTG 732
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
293 .....LeuPro 295
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
733 CCCGACGCAAGCGCTGCTGTCACATCGCGCCGTCGAAGGGA... 779
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
295 rGArgGlnAspArgValArg...AlaHisArgProArgProArgGlyArg 310
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
780 .....ATGACGCGCAACCAAGCCGACGATGACGCGCGCTGT 814
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
311 GlyGlyAlaAlaArgLeuAspArgAlaGlyGlyValGlyAlaArgPro 327
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
815 TCAACCGCA 824
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
327 ArgProArg 330
   ||| :|||:|||||:|||||:|||||:|||||:|||||:
```

seq\_name: p1r2:J00405

seq\_documentation\_block:  
 hypothetical 119.5K protein (uvra region) - Micrococcus luteus

N:Alternate names: ORF 1 protein  
 C:Species: Micrococcus luteus, Micrococcus lysodeikticus

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 04-Feb-2000

C:Accession: J00405

R:Shiota, S.; Nakayama, H.  
 Mol. Gen. Genet. 217, 332-340, 1989

A:Title: Micrococcus luteus homolog of the Escherichia coli uvra gene: identification

A:Reference number: S04781; MUID:89364717

A:Accession: J00405

A:Molecule type: DNA

A:Residues: 1-1106 <SHIT>

A:Cross-references: EMBL:X15867

A>Note: this reading frame extends between two stop codons and does not begin with a

A>Note: the gene encoding this protein overlaps uvra gene

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homo

alignment\_scores:  
 Quality: 145.00 Length: 366  
 Ratio: 1.007 Gaps: 19  
 Percent Similarity: 39.344 Percent Identity: 25.683

alignment\_block:  
 US-09-303-518D-571 x J00405 ..

Align seg 1/1 to: J00405 from: 1 to: 1106

```

18 CAGCGTGTTCCTCCCTTTCGAACGCG.....CA 46
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
```







seq\_name: p1r2:F82350

seq\_documentation\_block:

1 Lipid A biosynthesis (kds)-1 Lipid IVA acyltransferase VC0212 [Imported] - Vibrio cholerae

2 CDate: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

3 CAccession: F82350

4 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

5 charidson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.

6 Nature 406, 1477-483, 2000

7 A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

8 A:Reference number: A82035; MUID:20406833

9 A:Accession: F82350

10 A:Status: Preliminary

11 A:Molecule type: DNA

12 A:Residues: 1-325 <HE>

13 A:Cross-references: GB:AE004111; GB:AE003852; NID:9654614; PIDN:AAF93388.1; GSPDB:GN001

14 C:Experimental source: serogroup O1, strain N16961, biotype El Tor

C:Genetics:

A:Gene: VC0212

A:Map position: 1

```
alignment_scores:      193
      Quality:      137.50
      Ratio:      1.239
      Gaps:      6
Percent Similarity:  57.513
Percent Identity:   24.870
```

alignment\_block:  
nc-00-303-519D-571 x F82350

Align seg 1/1 to: F82350 from: 1 to: 325

[illegible]



```

181 ATGGCGGAGCGGGTTTGAACCCGACGACGACGAGCGTCAAGCCGTTT 230
    |||
62 LeuAlaLeuAla.....PheProGluLysThrPh 71
231 TGGGAAAGCGCAAAATGCGT..... 252
    |||
71 AspGluArgTyrLysIleAlaArgGlnSerLeuGlnHisLeuIleIer 88
253 .....TTGGAAGCTTGGC..... 264
    |||
88 hrLeuLeuGlnLeuLeuAlaIleGlnLeuValGlyAsnIleAspLys 104
265 .....CCGCGGCTTTTCAAAAAACC 284
    |||
105 LeuIleThrIleValThrSerSerArgAsnProLysGlyPheSerSerI 121
285 G.....GAAGACATGCAAAACATGTTCAAAAGCGGTACACG 319
    |||
121 uGluValIleSerAsnGlnAspLeuGlnGluThrPheLys..... 134
320 GCTGGGAAACGCTGCAGCAGCTTTGACAAAGGCGAAGGCGTCTTTC 369
    |||
135 .....AsnLeuGlnGlnLysGlnGlnLeuLeu 144
370 ATCAGCGCGCACATCGCAGCTACGATTGGCGGACGCTACATGACCA 419
    |||
145 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrIleThr 161
420 GAGAGCTCCGTTCCACCTGACCGCCATGACAAAGCCGCAAAATCAAG 469
    |||
161 sAsnTyrProGly.....IleAlaPheAlaLysAlaIleLysAsnGln 176
470 CGATGACAAATCATGACGCGGCGGCGGCGGCAAAAGCAAAACC 519
    |||
176 rGlnSerLysLysIlePheAlaLeuArgGluValPheLysGlyLysIle 192
520 GCGCCACCGCGCATACAGGGGTCAAAACATATCAGAGCCCTGCGCGC 569
    |||
193 ValProPheLys...AsnGlyIleGlnGlnGlyIleGluAlaLeuAsnG 208
570 GGGCGAGGACCATCATCTGCGCCGACACGTC..... 603
    |||
208 nGlyLysLeuValGlyIleValGlyAspGlnAlaLeuLeuMetSerT 225
604 .....CCTTCCGCGAGAAAGCGCGGCGTGTGGCGGATTTTTC 645
    |||
225 yThrTyrPro.....LeuPhe 230
646 GGCAAACCTGCATACCATGACACTGCGGCAAAATGCGACACGTCAA 695
    |||
231 GlySerProAlaPheThrThrThrSerProAlaLeuAlaTyrLysTh 247
696 AGCGCGAAACCCCTGTTTCTGCTGCGACAGCCCTGCCGAGGACAG 745
    |||
247 rGlyPheProValIleAlaValAsnValSerArg.....GlnAlaLysG 262
746 GCTTGTGTTGCATCCGCCGTCACAGGGGAATGAAAGCGCAACAA 795
    |||
262 yPhe.....GluValIlePro...SerAlaLysLeuTyrAlaAsnLys 275
796 GCC.....CAGCATGCGCGCGTTCACACCGCAAT..... 825
    |||
276 SerLeuProMetLysGlnSerValAlaIleLeuMetAspGlnMetMetG 292
826 ....ACCGAATATGGATACCGCGCTTTCCGACGCGAGTATCTGTTATGT 871
    |||
292 yPheLeuGlnLysGlyIleAlaSerGlnProGlnGlnTrpMetTrpIleH 309
872 ACAACCGCTATATA 885
309 IsLysArgTrpLys 313

```

```

seq_name: p12:B72119
seq_documentation_block:
acyltansiferase - Chlamydia pneumoniae (strain CWL029)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: B72119
R:Kallen, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: B72119
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <ARN>
A:Cross-references: GB:AE001596; GB:AE001363; NID:94376357; PIDN:ADD18251.1; PID:9437
A:Experimental source: strain CWL029
A:Genetics:
A:Gene: htrB

alignment_scores:
    Quality: 137.00      Length: 338
    Ratio: 0.867        Gaps: 15
    Percent Similarity: 46.746      Percent Identity: 23.964

alignment_block:
us-09-303-518d-571 x B72119 ..

Align seg 1/1 to: B72119 from: 1 to: 467

52 ATCCGTGTGACCGCCCTGCTCAAAATCCCTCCCTGCTGCTTTCG 101
    |||
12 rLeuGlnAlaProLeuTyrTyrLeuValSerGlyIleIleAlaLeu 28
102 TCTGCACAG.....CTGGAAACCGGCTCGGAC 130
    |||
28 sArgHisThrProArgSerPheLeuThrGlyLysGlyPheGlyP 45
131 ATCTGCGTTTACCTTTAAAGAGACGCGCGCATCTGCCAAT 180
    |||
45 hLeuAlaPheTyrIleIleSerAspTyrArgLysThrAlaLeuThrAsn 61
181 ATCCGCGACGCGGTTTGAACCCCGACGACGAGCGTCAAGCCGTTT 230
    |||
62 LeuAlaLeuAla.....PheProGluLysThrPh 71
231 TGGGAAAGCGCAAAATGCGT..... 252
    |||
71 AspGluArgTyrLysIleAlaArgGlnSerLeuGlnHisLeuIleIer 88
253 .....TTGGAAGCTTGGC..... 264
    |||
88 hrLeuLeuGlnLeuLeuAlaIleGlnLeuValGlyAsnIleAspLys 104
265 .....CCGCGGCTTTTCAAAAAACC 284
    |||
105 LeuIleThrIleValThrSerSerArgAsnProLysGlyPheSerSerI 121
285 G.....GAAGACATGCAAAACATGTTCAAAAGCGGTACACG 319
    |||
121 uGluValIleSerAsnGlnAspLeuGlnGluThrPheLys..... 134
320 GCTGGGAAACGCTGCAGCAGCTTTGACAAAGGCGGAGCGTCTTTC 369
    |||
135 .....AsnLeuGlnGlnLysGlnGlnLeuLeu 144
370 ATCAGCGCGCACATCGCAGCTACGATTGGCGGACGCTACATGACCA 419
    |||
145 PheCysGlyHisGlnAlaAsnTrpGluLeuProPheLeuTyrIleThr 161
420 GAGAGCTCCGTTCCACCTGACCGCCATGACAAAGCCGCAAAATCAAG 469
    |||

```



```

161 SASnTyProGly.....IleAlaPheAlaLysAlaLysAsnGln 176
470 CGATGACAAATATCATGCGAGCGGCGGCGGCAAGGCAAAAC 519
176 rglLeuSerLysLysIlePheAlaLeuArgLysValPheLysGlyLysIle 192
520 GCGCCACCGGCGATACAGGGGTCAAAACAAATCATCAAGCCCTCGCCG 569
193 ValProProls...AsnGlyIleGlnGlnIleLysIleGlnAlaLeuAsnGln 208
570 GGGCGAGCAGCAGCATCATCTCCCGACACGCTC..... 603
208 nglLysLeuValGlyIleValGlyAspGlnAlaLeuLeuMetSer 225
604 .....CTTCTCCGAGAGCGCGCGCTGTGGCGCATTTTTC 645
225 yTrnTyPro.....LeuPhe 230
646 GGCAGACCTGCATACACGACACGCGGCAAAATGGCACAGTCAA 695
231 GlySerProAlaPheTrnThrThrSerProAlaLeuAlaTyLysTh 247
696 AGGGGTGAACCCCTGTTTCTCTGCGAAGCGCTCGCCGACGACAG 745
247 rglPheProAlaIleAlaValAlaValSerArg.....GlnAlaLysG 262
746 GCTTCGTGTTGACATCCGCGCCGCGCAAGGGAATTGAACGCGCAAA 795
262 LysPhe.....GluValIlePro...SerAlaLysLeuTyAlaAsnLys 275
796 GCC.....CACGATGCGCGCGTGTCAACCCGCAAT..... 825
276 SerLeuProMetLysGlnSerValAlaIleLeuMetAspGlnMetGln 292
826 ....ACCGAATATGATGACCGCGCTTTCGACGACGATCATCTGTTATG 871
292 yPheLeuGlnLysGlyIleAlaSerGlnProGlnIntrPrMetTrpIle 309
872 ACAACCGCTATATAA 885
309 lAspArgTrpLys 313
seq_name: p1r2:D97552
seq_documentation_block:
lipid a biosynthesis lauroyl acyltransferase (htrb) rp718 [imported] - Agrobacterium tum
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: D97552
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Durillo, B.; Goldman,
A.; Liu, F.; Moilam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2333-2338, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: D97552
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 (KUR>)
A:Cross-references: GB:AE007869; PIDN:AAK87373.1; PID:G15156679; GSPDB:GN00169
A:Gene: AGR_C_2930
A:Map position: circular chromosome
alignment_scores:
Quality: 136.50 Length: 287
Ratio: 0.982 Gaps: 14
Percent Similarity: 48.432 Percent Identity: 22.648
alignment_block:
US-09-303-518d-571 x D97552 ..
Align seg 1/1 to: D97552 from: 1 to: 312

```

```

190 GCGGGTTGACCCCGACACGCGAGCGTCAAGCCGTTTTCGGAAC 239
41 AlagIleArgAlaThrAspArgLeuAlaArgLeuIleGlyProLysTh 57
240 GGCAGAAATGCGGTTTG.....GACTTGCCCGCGGTTTTCAAA 280
57 rglAlaGlnLysLeuMetLeuTyTrnLeuAlaAlaGlnAlaPheProGln 74
281 MACCGAAGACATCGAACAATGTTCAAGCGGTACACGCTG..... 324
74 yTrnGlnGln.....GluArgLeuAlaIleAlaMetAspSerTrpAlaAsn 89
325 .....GACACGTC..... 333
90 MetGlyArgLeuAlaIleGlnTyValPheLeuAspArgLeuPheAsp 106
333 ..... 333
106 eaSPProGlnLysAsnGlnProGlyArgIleGlnValGlnGlyTrnSer 123
334 ....CACAGGCTTTGACAGAGGCGGAGGCGCTGTTTCATCAGCGG 378
123 hrPheLeuGlnLeuArgAspAsnProArgProPheIleValPheThrAla 139
379 CACATCGGAGCTACGATTTGGCGGCGACCTACATCAGCCAGCAGCTTC 428
140 HisSerGlyAsnPheGlnLeu.....LeuTr 148
429 G.....TTCACCTGACCGCATGTACAAGCGC 457
148 oValAlaGlnSerAlaPheGlyLeuAspValThrValLeuPheArgPro 165
458 CCAAAATCAAGGATA...GACAAATCATCAGCGGCGAGGTGCGC 504
165 roAsnAsnProTyValAlaAspLysValPheAsnPheArgLysGlnArg 181
505 GGCAGAGCAAAACCGCGCCGACGCGCATACAAAGGGGTCAAAATCAT 554
182 ...MetGlyAsnLeuValProSerThrAlaGlySerPheAlaLeuAl 197
555 CAAGGCGCTGCGCGCGGCGGAGCAACCATCTCTGCGGACGACGTC 604
197 aArgGlnLeuGlnLysGlyGlyValGlyValLeuValAspGln.... 212
605 CTTCTCCGAGAGAGCGCGCGCTGTGGCGGATTTTTCGGAACCT 654
213 .....LysPheGlyLysGlyLeuTrnThrLysPhePheGlyLeuGln 226
655 GCATACACCATGACATGCGCGCAAAATTCGACACGTCMAAGGGGTGA 704
227 ValArgThrAsnProLeuLeuAlaLysLeuValArg..... 238
705 AACCGCTTTTCTG.....TCGACGCGCTGCGG 736
239 ....GlnPheAsnCysAspValTyProAlaArgLysIleArgLeuPro 254
737 ACGGACAAAGCTTCGTTGACATCCGCGC.....GTCAAAGG 777
254 sPAsnArg...TyTrnGlnGlnIleGlnProLysValGlnIleProArg 269
778 GAATTGAAGCGCAACAA.....GCCACGATGCGCGCGTTTCAACG 821
270 AspGlnLysGlyAsnValAspIleGlnAlaThrAlaGlnLeuLeuAsnAs 286
822 CAATACCAATATTGGATACCGCGTTTTCGACGACGATGTCTTTATGT 871
286 pLysValGlnSerTrpValArgGlnTyProGlyGlnIntrPrLeuTrpTy 303
872 ACAACCGCTAT 882
303 lAspArgTrp 306

```







```

269 spleuAlaAspAlaAspAlaTyrIleAlaArgMetAsnIleu 285
      ||| : : : : : ||| : : : : :
826 ACCGAATATGATACGGCTTTCCGACGACATCTGTTATGACAA 875
      : : : : : : : : : : : : : : :
286 ValGluLeuLeuValIleProAsnProGluInIlyrThrIleLeu 302
      : : : : : : : : : : : : : : :
876 CCGCTATAAACG 888
      : : : : :
302 spleuLeuIlyrThr 306

seq_name: p1r2:T14180

seq_documentation_block:
exit protein - Mycobacterium smegmatis
C:Species: Mycobacterium smegmatis
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C:Accession: T14180; T14163
R:Zhu, W.M.; Arceneaux, J.E.L.; Beggs, M.L.; Byers, B.R.; Eisenach, K.D.; Lundrigan, M.D
Mol. Microbiol. 29, 629-639, 1998
A:Title: Exochelin genes in Mycobacterium smegmatis: identification of an ABC transporter
A:Reference number: Z17906; MUID:98385832
A:Accession: T14180
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1122 <ZHU>
A:Cross-references: EMBL:AF034152; NID:g3421057; PID:g3421058; PIDN:AAC32046.1
R:Yu, S.; Fiss, E.; Jacobs Jr., W.R.
J. Bacteriol. 180, 4676-4685, 1998
A:Title: Analysis of the exochelin locus in mycobacterium smegmatis: biosynthesis genes
A:Reference number: Z17898; MUID:98389687
A:Accession: T14163
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 534-1122 <YUS>
A:Cross-references: EMBL:AF027770; NID:g3560502; PID:g3560505; PIDN:AAC82548.1
C:Genetics:
A:Gene: exit
C:Superfamily: Mycobacterium tuberculosis probable ABC transporter RV0194; ATP-binding c

alignment_scores:
Quality: 130.00 Length: 337
Ratio: 0.828 Gaps: 18
Percent Similarity: 46.588 Percent Identity: 26.409

alignment_block:
US-09-303-518D-571 x T14180 ..
Align seg 1/1 to: T14180 from: 1 to: 1122

3 GTTTCCTTACATTCAGGCTGTTCCCTTGGCAACGCCGACGACACA 52
||| ||| : : : : : ||| : : : : :
135 ValValIleGlyLeuGlnLeuThrAlaProMetIleAlaArgValGlu 151
      : : : : : : : : : : : : : : :
53 TCCTGTTGACCGCCCTGCAATAGCCTGCTCCCTGCTGCGCTTCCTGT 102
      : : : : : : : : : : : : : : :
151 u.....AspGlnGlnAlaGlnIleIly.....A 159
      : : : : :
103 CTGCACACCGCTGGAAACCGGCTCGACATCTGGCTTTACCTTTAA 152
||| ||| : : : : : ||| : : : : :
159 rGAlaThrAlaLeuAlaThrAspLeuIleSerGlyGlnArgProLeuGln 175
      : : : : : : : : : : : : : : :
153 GGAAGACCGCGCGCATCGTCCCAATATGCGGACGCGGGTTGAAC 202
||| : : : : : : : : : : :
176 GlyAsnAlaArgAlaGlyGlnArgIleAlaLeuProCysArgGluPr 192
      : : : : : : : : : : :
203 CCGACACGACGAC..... 215
      : : : : :
192 oSerThrAlaAspArgAspValAlaCysArgAlaHisProGluHisProL 209
      : : : : :
216 ..GTCGAAAGCCGTTTTCGGAAACGCGCAAAATGCGGTTGAAC... 260
      : : : : :

```

```

209 euGIyArg.....CysGIyGIyArg.....GIyThrArg 219
261 .....TGCCCCCGCGCTTT 274
220 ArgAspGIyGIyArgGlyArgGlyLeuLeuArgAspProArg...Ph 235
275 TCAGAAAAACCGA.....AGAC 291
      : : : : :
235 eAspAspGIyGIyProValAspHisArgAspArgProArgProValProH 252
292 ATCGAAACATGTTCAAGCGGTACACG.....CTGGACACGTGCA 335
||| : : : : : ||| : : : : :
252 IsArgThrValLeuAlaIleGlyArgAlaGlnLeuIleSerArgGly 268
336 GCAGCTTTGGACAAAGCGGAAAGGCTGTGTCATACGCGCGACATCG 385
      : : : : :
269 ProGlyValGIyGlnProArgIleSerArgProGluArgProAspProR 285
386 GCAGCTACGATTT...GGCGGACGCTACATACGCCAGCGCTCCGTTG 432
285 oleuProArgLeuArgGlyGIyProProArgArgArgGlyProProValG 302
433 CACCTGACCGCCATGTACAGCCGCGAA.....AATCAAGC 470
||| : : : : : ||| : : : : :
302 LyProGlnArgAsnThrArgGlySerGlnLyProValAlaArgArgArg 318
471 GATAGCAAAATCATGACAGCGCGGACGGTGGCGCAAGCGAAACCG 520
||| : : : : : ||| : : : : :
319 AlaArgArgIleArgArgGlyThrHisGluArgArgProArgCysArgAr 335
521 CCCCCACCGGCAT.....ACAGGGGTCAACAAATC 552
||| : : : : : ||| : : : : :
335 gThrHisArgSerAlaLeuArgIleSerAlaSerArgGlyGlnArgHisG 352
553 ATCAAGGC.....CTGCGCGCGCGGACGAGC 578
||| : : : : : ||| : : : : :
352 LyAlaGlyArgThrProAlaProArgProAspProAlaGlu...ArgAla 367
579 AACCATCATCTGCGCCGACACGCTCCTTCCGCGAGAGCGCGCG 628
      : : : : :
368 ProArgThrProAlaArgArgThrProProGlnLeuValAlaGlnArgAs 384
629 TGTCGGCGGATTTTTCGCAACCTGCATACACCATACGACGCGCGCA 678
||| : : : : : ||| : : : : :
384 PalAlaIlyAsn.....GlnSerGluCysHisArgAsnArgSerAlaAla 399
679 AAATGGCACACGTCGAAAGCGGTGAAACCTGTTTTCGCTGCGAAGC 728
      : : : : :
399 rGThrGlyLeuArgArgArgThrThrGlyVal..... 410
729 CCGCGCGAGGACAGAGCTTGTGTCACATCCGCCGCTCCAGGG 778
||| : : : : : ||| : : : : :
411 ..ValArgArgArgArgArgArgProAla...ProArgArgSerArgPr 425
779 AATTGAAGCGCAACAAAGCCGACGATGCGCGCTGTCACCGCAATAC 828
||| : : : : : ||| : : : : :
425 ocYAsArgArg.....GlyAlaArgCysGlnProValAlaArgThrThrP 440
829 GAATATTGGAT 839
      : : : : :
440 roThrLeuAsp 443

seq_name: p1r2:A35938

seq_documentation_block:
profilaggrin - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 14-Dec-1990 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: A35938
R:Gan, S.O.; McBride, O.W.; Idler, W.W.; Markova, N.; Steinert, P.M.
Biochemistry 29, 9432-9440, 1990.
A:Title: Organization, structure, and polymorphisms of the human profilaggrin gene.
A:Reference number: A35938; MUID:91064347

```

A:Accession: A35938  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-2248 <G>  
 A:Cross-References: GB:J029929  
 C:Genetics:  
 A:Gene: GDB: FLG  
 A:Cross-References: GDB:119912; OMTM:135940  
 A:Map position: 1q21-1q21  
 C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology  
 C:Keywords: EF hand; epidermis; polymorphism; tandem repeat  
 F:246-569/Region: filaggrin repeat  
 F:570-893/Region: filaggrin repeat  
 F:1074-1397/Region: filaggrin repeat  
 F:1573-1896/Region: filaggrin repeat

alignment\_scores:  
 Quality: 127.50 Length: 236  
 Ratio: 1.081 Gaps: 10  
 Percent Similarity: 50.000 Percent Identity: 25.000

alignment\_block:  
 US-09-303-518D-571 x A35938 ..

Align seg 1/1 to: A35938 from: 1 to: 2248

```

198 GAACCCGACACGACGACGTCACACCGCTTTTGGCGAAGCGCAAAAT 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
725 Glnsppatgaspsergluglglinsergln.....aspsergluar 738
248 GCGGTTTGACTGTGCCCGCGCTTTTCAAAAAACGGAAGACATGAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
738 ghiserclyseralaseralg.....Asnlsarlg 749
298 ACAATGTTCAAGCGGTACACGCTGGACACGTCGACGCGCTTGA 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
749 lyseralglinslucinseralgspglserarghnsproglserhns 765
348 CAAGGCGAAGGCGTGTTCATCAGCCGCA..... 380
766 Aspgluaspargalaglghisarglinseralaspserserglnse 782
381 .....CATCGCAGCTA.....CGATTGGCGGACCGCTAC 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
782 rcllythrghisrghnthrserserargrghlnlalaasergrg 799
412 ATCAGCCAGCAGCTCCGTCACCTGACCGCATGTACAAGCGCGCAA 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
799 lnglglnlalargserargalaglasparghnsghlserghlsgln 815
462 AATCAAGCGATAGACAAATCATCAGCG.....GGCAGGCTGC 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
816 GlnserAlaspserserarghnsghlserghlglargghlglalase 832
503 GCGCGCAAGGCAAAACCGCGCCGCGCATCAGGGGT..... 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
832 rserlavalargasparglghlserargghlserghlserghlnlals 849
543 .....CAACAATCATCA.....GGCCT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
849 ertspglnlucghlserghlaspseraspserghlserghlserghl 865
564 GCGCGGGGCGAGCAACCAT.....CATCTGCCCGACACGCTGC 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
866 Glnargghlglalghlserghlsglninsrghlsglnlserghl 882
605 CTCTCCGAGGAGGCGCGGCGGTGGCGGATTTTTCGCAACACT 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
882 yarghlserghlghlserghlghlserghlserghlserghlserghl 899
655 GCATACCATGACACTGCGCGCAAAATTTGC.....ACA 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

899 ertghnslglinslucinseralahnsghlserghlserghlserghl 915  
 690 CGTCAAGCGCTGAACCCGCTTTTCTGCTGGACGCGCTGCCGACG 739  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 916 Glnargghlglinslucinslucinslucinslucinslucinsluc 932  
 740 GACAAGCGCTGCTGT.....GCACATCCGCCCGCTC 771  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 932 sserlaserghlglinslucinslucinslucinslucinslucinsl 949  
 772 CAAGCGCA 779  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 949 rghlserghl 951  
 seq\_name: p1r2:E90949  
 seq\_documentation\_block:  
 heat shock protein msbB [imported] - Escherichia coli (strain O157:H7, substrain R1MD  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C:Accession: E90949  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shindagawa, H.  
 DNA Res. 8: 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
 A:Reference number: A95629; MUID:21156231; PMID:11258796  
 A:Accession: E90949  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <HAY>  
 A:Cross-References: GB:BA000007; PID:BA835988.1; PID:q13362033; GSPDB:GND0154  
 A:Experimental source: strain O157:H7, substrain R1MD 0509952  
 C:Genetics:  
 A:Gene: Ecs2565

alignment\_scores:  
 Quality: 127.00 Length: 274  
 Ratio: 0.914 Gaps: 11  
 Percent Similarity: 50.730 Percent Identity: 24.088

alignment\_block:  
 US-09-303-518D-571 x E90949 ..

Align seg 1/1 to: E90949 from: 1 to: 323

```

121 CGGCTGGACATCGCGCTTTACCTTTAAGAGACGCGCGCAT 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 Argleucllyargllealaglghlserghlserghlserghlserghl 66
171 CGTGGCCATATG.....CGGC 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66 aleuclleasnlsrserleuclserghlserghlserghlserghlser 83
188 AGCGGGTTTGAACCCCGACACGACGTCACACCGCTTTTGGCGAA 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 luala.....llevalaspghlserghlserghlserghlserghl 92
238 ACGCAAAATCGGCTTGAATTCGCGCGCGCTTTTCAAAAACCGCA 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
93 Alaproglnalmetalaalmetalaalmetalaalmetalaalmetala 109
109 ulyslileglproargvalasp...trpglghlserghlserghlser 125
338 AGGCTTTGGACAGGCGGAGGCGGTGTTCATCAGCGCGCATCGCG 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 lmetargargasnslglvalilghlserghlserghlserghlserghl 141
388 AGCTACGATTTGGCGGCGCGCATCAGCGCGCGCGCTTCGTCACCT 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 Alavalaspilserghlserghlserghlserghlserghlserghl 157

```

```

438 GACCGCATGTACAGCCGCGAAATCAAGCATAGACAAATCATGCG 487
      ::::::::::::::::::::
157 tAlaAlaMetPheNHisAsnGlnGlyAsnProValPheAspIyValITrPA 174
      ::::::::::::::::::::
488 AGCGGGGAGGGGTGGCGGCAAGGCAAAACCGCCCGCGCATACAA 537
      ::::::::::::::::::::
174 sTrhValaIrgArGArGpHeGlyGArGLeuNHisAlaIrgAsn...Asp 189
      ::::::::::::::::::::
538 GGGGTCAAAATCATCAAGCGCCCTGGCGGCGGCGGAGCAACATCAT 587
      ::::::::::::::::::::
190 GlyIleIysProPheIleGlnSerValaIrgGlnIleIyTrpGlyTyTy 206
      ::::::::::::::::::::
588 CCTGCC.....GACCACTCCCTTCTCCGCAAGGCGCGCGCGCTGT 631
      ::::::::::::::::::::
206 rLeuProAspGlnAspHisGlyProGlnHisSerGlu.....P 219
      ::::::::::::::::::::
632 GGGCGGATTTTTCGGCAACCTGCATACACATGACATGCGCGCAAAA 681
      ::::::::::::::::::::
219 heValAspPhePheAlaThrTyIysAlaThrLeuProAlaIleGlyArg 235
      ::::::::::::::::::::
682 TTGGCACACGTC.....AAAGCGGTGMAAACCTGTTTTCTGCTCGA 725
      ::::::::::::::::::::
236 LeuMetIysValIysAspAlaIrgValIValProLeuPhe..... 248
      ::::::::::::::::::::
726 ACGGCTGGCGGACGCA.....GGCTTCGTGTTGCACATCGCGCGG 769
      ::::::::::::::::::::
249 .ProIleIyAspGlyIysThrNHisArgLeuThrIleGlnValaIrgPro 265
      ::::::::::::::::::::
770 TCCAGGGAATTTGAACGGCAACAAACCCACAT...GCCGCGGTTC 816
      ::::::::::::::::::::
265 roMetAspAspLeuLeuGlnAlaAspAspHisThrIleAlaIrgArGmet 281
      ::::::::::::::::::::
817 AACCGCAATACCGAATTTGGATACGGCGCTTTCGACGCAATATCTGT 866
      ::::::::::::::::::::
282 AsnGlnIleValGlnIlePheValIglyProArgProGlnIleIyTrhTr 298
      ::::::::::::::::::::
867 TATGTACAAACCGGTATTAACG 888
298 pIleLeuIysLeuLeuIysThr 305

```

seq\_name: pIir2:A85798

seq\_documentation\_block:

suppressor of htrb, heat shock protein [imported] - Escherichia coli (strain O157:H7, su  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: A85798  
 R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: A85798  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-923 <STO>  
 A:Cross-references: GB:AE005174; NID:912515911; PIDN:AAG56845.1; GSPDB:GN00145; UWGP:Z29  
 C:Genetics:  
 A:Gene: msbB

alignment\_scores:

Quality:	127.00	Length:	274
Ratio:	0.914	Gaps:	11
Percent Similarity:	50.730	Percent Identity:	24.088

alignment\_block:

US-09-303-518d-571 x A85798 ..

Align seg 1/1 to: A85798 from: 1 to: 323

```

121 CGGCTCGACATGTGGCGTTTACCTTTTAAAGAAAGCCGCGGCAT 170
      ::::::::::::::::::::
50 ArgLeuIyArGlyIleAlaIyArGLeuGlyIysSerSerIyArGArG 66
      ::::::::::::::::::::
171 CGTCCCAATATG.....CGGC 187
      ::::::::::::::::::::
66 aleuIleAsnLeuSerLeuCysPheProGluArGSerGluAlaIyArG 83
      ::::::::::::::::::::
188 AGCGGGTTTGAACCCCGACACGACAGCGTCAAGCCGTTTTCGGGAA 237
      ::::::::::::::::::::
83 lAla.....IleValAspIleMetPheAlaThr 92
      ::::::::::::::::::::
238 ACGGCAAAATGCGGTTTGGAACTTGGCCCGGCTTTTAAAAAACGGA 287
      ::::::::::::::::::::
93 AlaProGlnAlaMetAlaMetMetAlaIyLeuAlaIleArgGlyProG 109
      ::::::::::::::::::::
288 AGACATCGAAATCATGTCMAAGCGGTACAGCGCTGGCAACGTCACG 337
      ::::::::::::::::::::
109 uIysIleGlnProArGValAsp...TrpIleIyLeuGlnIleIleGlnG 125
      ::::::::::::::::::::
338 AGCTTTGGACAAAGCGCAAGCGCTGCTGTATCAAGCGCGCATCGGC 387
      ::::::::::::::::::::
125 lMetArGArGAsnGlnIyValIlePheLeuValProHisGlyTrp 141
      ::::::::::::::::::::
388 AGCTACGATTTGGCGGACGCTACATCAAGCAAGCTTCCTTCACCT 437
      ::::::::::::::::::::
142 AlaValAspIleProAlaMetLeuMetAlaSerGln...GlyGlnIyMe 157
      ::::::::::::::::::::
438 GACCGCATGTACAGCCGCGAAATCAAGCATAGACAAATCATGCG 487
      ::::::::::::::::::::
157 tAlaAlaMetPheNHisAsnGlnGlyAsnProValPheAspIyValITrPA 174
      ::::::::::::::::::::
488 AGCGGGGAGGGGTGGCGGCAAGGCAAAACCGCCCGCGCATACAA 537
      ::::::::::::::::::::
174 sTrhValaIrgArGArGpHeGlyGArGLeuNHisAlaIrgAsn...Asp 189
      ::::::::::::::::::::
538 GGGGTCAAAATCATCAAGCGCCCTGGCGGCGGAGCAACATCAT 587
      ::::::::::::::::::::
190 GlyIleIysProPheIleGlnSerValaIrgGlnIleIyTrpGlyTyTy 206
      ::::::::::::::::::::
588 CCTGCC.....GACCACTCCCTTCTCCGCAAGGCGCGCGCGCTGT 631
      ::::::::::::::::::::
206 rLeuProAspGlnAspHisGlyProGlnHisSerGlu.....P 219
      ::::::::::::::::::::
632 GGGCGGATTTTTCGGCAACCTGCATACACATGACATGCGCGCAAAA 681
      ::::::::::::::::::::
219 heValAspPhePheAlaThrTyIysAlaThrLeuProAlaIleGlyArg 235
      ::::::::::::::::::::
682 TTGGCACACGTC.....AAAGCGGTGMAAACCTGTTTTCTGCTCGA 725
      ::::::::::::::::::::
236 LeuMetIysValIysAspAlaIrgValIValProLeuPhe..... 248
      ::::::::::::::::::::
726 ACGGCTGGCGGACGCA.....GGCTTCGTGTTGCACATCGCGCGG 769
      ::::::::::::::::::::
249 .ProIleIyAspGlyIysThrNHisArgLeuThrIleGlnValaIrgPro 265
      ::::::::::::::::::::
770 TCCAGGGAATTTGAACGGCAACAAACCCACAT...GCCGCGGTTC 816
      ::::::::::::::::::::
265 roMetAspAspLeuLeuGlnAlaAspAspHisThrIleAlaIrgArGmet 281
      ::::::::::::::::::::
817 AACCGCAATACCGAATTTGGATACGGCGCTTTCGACGCAATATCTGT 866
      ::::::::::::::::::::
282 AsnGlnIleValGlnIlePheValIglyProArgProGlnIleIyTrhTr 298
      ::::::::::::::::::::
867 TATGTACAAACCGGTATTAACG 888
298 pIleLeuIysLeuLeuIysThr 305

```

seq\_name: pIir2:T43023

seq\_documentation\_block:

ceoA protein - Burkholderia cepacia  
 C:Species: Burkholderia cepacia





```

334 CAGCAGCGCTTTGGACAGAGGCGGCTGCTGTTTCATCAGCGGCACAT 383
      |||||
142 rlysaargalathrthrargasn.....Prothrlyargthrpt 156
      |||||
384 CG.....GCAGCTACGATTTGGCGGAGCGGTACATCAGCGGAGC 424
      |||||
156 hrThrlyargalathrThrlysaargalathrThrlysaargalathrThr 172
      |||||
425 TTCGCTTCAGCTGACCGCATGACAAAGCGCGGAAATCAACAGCATA 474
      |||||
173 ArgArgasnProthrlysaarglystThrlyargthrThrlysaargal 189
      |||||
475 GACAAATATCATGACAGCGGCGGAGGTGCGGCAAAAGCAAAACCGCGC 524
      |||||
189 athrlysthThrlysaargala.....ThrThrlysaarg 201
      |||||
525 CACGCGCTACAGGGGCAAAATCAATCAACAGCGCTGCGCGGCGG 574
      |||||
201 laprothrlysaargalathrThrlysaargala...Prothrlysaargval 216
      |||||
575 .....AGCAACCATCATCTCTGCGGACGACGACGCTTCTCGCAG 615
      |||||
217 ThrThrlysaargalaprothrlysaargalathrThrlysaargalaproth 233
      |||||
616 GAAGCGCGCGCGCTGCGGCGGATTTTTCGCAAACTGCATACACAT 665
      |||||
233 rlysaargalathrThrlysaargala..... 241
      |||||
666 GACACTGCGCGCAAAATGCGACACGCGCAAGCGGTGAAACCGCTTTT 715
      |||||
242 .....Prothrlysaargala..... 246
      |||||
716 TCTGCTCGCAAGCGCTGCGGACGACAGCTTCTGTTGCATCCGC 765
      |||||
247 ...ThrThrlysaargalaprothrlysaargalathr...Thrlysaargal 261
      |||||
766 CCGCTCCAGCGGATTTGAACG.....GCACAAAGCGCGACGATGC 806
      |||||
261 aProthrlysaargalathrThrlysaargalathrAlaargProthrSertL 278
      |||||
807 CGCGCTGT 814
      |||||
278 yseProcys 280
      |||||
seq_name: p1r2:A42608

seq_documentation_block:
(kdo)2-(lauroyl)-lipid IV4 acyltransferase (EC 2.3.1.-) - Escherichia coli
M:Alternate names: membrane-bound lytic transglycosylase
C:Species: Escherichia coli
C:Date: 04-Mar-1993 #sequence revision 18-Nov-1994 #text-change 21-Jul-2000
C:Accession: A42608; 154979; 664947
R:Karow, M.; Georgopoulos, C.
J. Bacteriol. 174, 702-710, 1992
A:Title: Isolation and characterization of the Escherichia coli msbB gene, a multicopy s
A:Reference number: A42608; MUID:92121107
A:Accession: A42608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <R>
A:Cross-references: GB:M87660; NID:q146876; PIDN:AAA96706.1; PID:q146877
A:Experimental source: strain K-12, substrain W3110
R:Engel, H.; Smink, A.J.; Van Wijngaarden, L.; Keck, W.
J. Bacteriol. 174, 6394-6403, 1992
A:Title: Morein-metabolizing enzymes from Escherichia coli: existence of a second lytic
A:Reference number: 154979; MUID:93015688
A:Accession: 154979
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-323 <R>
A:Cross-references: GB:M87660; NID:q146876; PIDN:AAA96706.1; PID:q146877
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cd

```

```

A: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: G64947
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-323 <R>
A:Cross-references: GB:AE000279; GB:U00096; NID:q1788154; PIDN:AACT4925.1; PID:q17881
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: msbB; mlt
C:Function:
A:Description: transfers myristate or laurate, activated on ACP, to (kdo)2-(lauroyl)-
A:Pathway: lipopolysaccharide core biosynthesis
A:Note: decanoyl, palmitoyl, palmitoleyl and (R)-3-hydroxy-myristoyl-ACP are poor acy
C:Keywords: acyltransferase; lipopolysaccharide core biosynthesis; transmembrane prot
F:23-43/Domain: transmembrane #status predicted <TM2>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:133-153/Domain: transmembrane #status predicted <TM3>

```

## alignment\_scores:

Quality:	Ratio:	Length:	Gaps:
125.00	0.899	274	11
Percent Similarity:	50.730	Percent Identity:	24.088

## alignment\_block:

US-09-303-518D-571 x A42608

Align seg 1/1 to: A42608 from: 1 to: 323

```

121 CGCGCTGACATCTGCGGTTTACCTTTTAAGAGACCGCGCGCAT 170
      |||||
50 ArgleuglyArgpheaIaClyArgleuglyLysSerSerArgArgAl 66
      |||||
171 CGTGGCAATATG.....CGCG 187
      |||||
66 alauIleasnSerLeuCySPheProGluArgSerGluAlaGluArgG 83
      |||||
188 AGCGGGTTTGAACCCGACGACGAGCGGTCAAGCGCTTTTTCGGGA 237
      |||||
83 luAla.....IleValAspGluMetPheAlaThr 92
      |||||
238 ACGGCAAAATGCGGTTTGAATTCGCGCGCGCTTTTCAAAAACCGGA 287
      |||||
93 AlaProGluAlaMetAlaMetAlaGluAlaIleArgGlyProG1 109
      |||||
288 AGACATCGAAGCATGTTCAAGCGGTACAGCGGTGGGAGACGAGGAGC 337
      |||||
109 uLysIleGlnProAlaArgValAsp...TyrGlnGlyLeuGluIleIleG 125
      |||||
338 AGCGTTTGGCAAGGCGGAGGCGGTGTTTCATCAGCGCGCATCGCGC 387
      |||||
125 luMetArgArgAsnAsnGluLysValIlePheLeuValProIleGlyTyr 141
      |||||
388 AGCTACGATTTGGCGGAGCGGTACATCAGCGAGCGCTTCGTTCCACT 437
      |||||
142 AlaValAspIleProAlaMetLeuMetAlaSerGln...GlyGlnLysSe 157
      |||||
438 GACCGCCATGTACAGCCGCGGAAATCAAGCGATGACAAATCATGCG 487
      |||||
157 talAlaMetPheIleAsnGlnGlnLysAsnProValPheAspTyrValTyr 174
      |||||
488 AGCGCGGAGGGTGGCGGCAAGGCAAAACCGCGCCACCGCATACAA 537
      |||||
174 snThrValArgArgPheGlyArgGluuHisAlaArgAsn...Asp 189
      |||||
538 GGGGTCAAAATCATCAAGCGCGGTGCGGCGGAGGAGGAGCATCAT 587
      |||||
190 GlyIleLysProPheIleGlnSerValArgGlnGlyTyrTyrGlyTyr 206
      |||||
588 CCGCGCC.....GACCACTGCTTCTCGCGAGGAGGCGGCGGTGT 631

```

```

206  rleprroaerglnlnarhnlsglyrproglinhlsrglu..... P 219
      ||||| ||||| |||::: ::::|
632  GGGGGGATTTTTTGGGCAAACTGTCATACACCATGACACAGCGGGGCAAA 681
      ::::| ||||| ||||| ::::| ||||| ::::|
219  hevalasprhealatlathrlyrlysalatlhrleuproalleglyatg 235
      ::::| ||||| ::::|
682  TTGGCAACACTG.....AAAGCGTGAAACCGCTGTTCCTGTCGCA 725
      ||||| ||||| |||||
236  LeumetylsalylsargylalargylalProleuphe..... 248
      |||||
726  ACGCGTCGCCGACGAGCAA.....GCGTCGTGTGTCACATCGCGCCG 769
      ::::| ||||| ::::|
249  .rpolletyraspolylulrlythlsatrgleumhrlleglinalrpropr 265
      ::::| ||||| ::::|
770  TCGCAAGGGGAATTGAACGGCAACAAAGCCACGAT...GCCGCGGTGTC 816
      ::::| ||||| ::::|
265  romklsarshryleuglualaaarhnlsthrllealargdymet 281
      |||||
817  AACGCAATACCGAATATGTGATACGCGCTTTCCGACGCGATCTGTT 866
      ||||| ||||| ::::| ||||| ::::|
282  Asncluglualgluileerhevalglrproargprogluglnlyrthtr 298
      ::::| ||||| ::::|
867  TATGTACACACCGCATTAACG 888
      ::::| |||||
298  rleleuylsleuileuylsthr 305

```